

UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



Doble Grado en Ingeniería Informática y Matemáticas

TRABAJO FIN DE GRADO

Evaluación Empírica de los Mecanismos de Usabilidad de Preferencias, Retroalimentación de Progreso y Abortar Operación en un Entorno Web

Autora: Herea Bouza Leirós
Tutora: Silvia Teresita Acuña Castillo

Enero 2016

Evaluación Empírica de los Mecanismos de Usabilidad de Preferencias, Retroalimentación de Progreso y Abortar Operación en un Entorno Web

AUTORA: Herea Bouza Leirós
TUTORA: Silvia Teresita Acuña Castillo

Grupo de Investigación de Herramientas Interactivas Avanzadas (GHIA)
Departamento de Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Enero 2016

*A Ana y Juan
por no dudar nunca y*

*A J. Juan
que aunque no pueda leerlo está orgulloso.*

Agradecimientos

Agradezco a todas las personas que han contribuido a alcanzar esta importante meta. En primer lugar, a mi tutora Silvia que con su amplia experiencia, conocimientos, consejos, disponibilidad y paciencia me ha ayudado a aprender durante este proceso.

Agradecer a Sira Vegas su ayuda, colaboración y tiempo. También quiero acordarme de Marcelo que sin conocerme siempre ha estado dispuesto a echar una mano así como de todo el Grupo de Investigación en Ingeniería del Software Empírica (GrISE).

Asimismo quiero dejar constancia de mi gratitud al numeroso grupo de estudiantes y profesores del Instituto de Rodeira y de la Universidad Autónoma de Madrid que colaboraron en la realización de este trabajo de una forma generosa y desinteresada.

Gracias a mi familia y amigos por ser el apoyo incondicional.

A todos muchas gracias.

Resumen

Cada día está más presente la importancia de la usabilidad en cualquier tipo de aplicación software y, concretamente, en aplicaciones con alta carga interactiva como una aplicación web. Sin embargo, no existen estudios experimentales que obtengan evidencias empíricas de hasta qué punto afecta la implementación de distintos mecanismos de usabilidad en la eficiencia y satisfacción de los usuarios finales de la aplicación.

Esta investigación nace con el objetivo de responder a esta pregunta. En concreto, se ha realizado un estudio empírico para determinar cómo afecta a la eficiencia y la satisfacción del usuario tener los mecanismos de usabilidad *Abortar Operación*, *Retroalimentación de Progreso* o *Preferencias* presentes o ausentes en una aplicación web de tienda, que gestione la comercialización online de artículos.

Para llevar a cabo el estudio se plantea un experimento verdadero que responde a un diseño *between subjects* donde existen tres factores de interés. El primero, es la presencia/ausencia del mecanismo de usabilidad que se está estudiando. El segundo, es el grupo. Existen cuatro grupos {A, B, C, D}. En los grupos A, B y C sólo el mecanismo *Preferencias*, *Abortar Operación* y *Retroalimentación de Progreso* está presente, respectivamente, mientras que en el grupo D los tres mecanismos de usabilidad están presentes. Es decir, en el 50% de los casos el mecanismo estará activo y en el 50% restante estará inactivo. El último factor es el orden de realización de las tareas diseñadas e implementadas para cada mecanismo, que representa las permutaciones de las tres tareas.

Una vez diseñado el experimento y creada la aplicación sobre la que éste es llevado a cabo se realizan varias sesiones con 100 sujetos experimentales. Cada uno debe realizar las tres tareas de las que consta el experimento, una tarea por cada uno de los mecanismos de usabilidad. Mientras los sujetos realizan el experimento, por una parte, se mide automáticamente el tiempo que éstos tardan en realizar cada una de las tareas y el número de clicks que utilizan. Ambas métricas permiten medir la eficiencia de los participantes en la realización de cada una de las tareas. Por otra parte, se solicita después de cada tarea que respondan a un cuestionario que ayuda a medir la satisfacción de cada usuario.

Tras la realización del experimento se plantea un estudio estadístico que se inicia analizando los estadísticos descriptivos básicos (media, mediana, desviación típica, coeficiente de asimetría, coeficiente de Curtosis, etc.) y los diagramas de caja para las distintas combinaciones de factores. En este primer análisis se detectan valores imposibles – datos que demuestran que el sujeto realizó incorrectamente la tarea, no son los valores atípicos o *outliers* – que son eliminados de la muestra.

Con la muestra refinada y tras verificar las hipótesis de normalidad de residuos y homogeneidad de las varianzas se plantea un test ANOVA para cada una de las variables respuesta (eficiencia y satisfacción) y para cada uno de los mecanismos de usabilidad estudiados (*Abortar Operación*, *Retroalimentación de Progreso* y *Preferencias*).

En el caso de la variable respuesta *satisfacción* los resultados son concluyentes: en los tres casos existe diferencia significativa entre la ausencia o presencia del mecanismo respecto a la satisfacción final de los sujetos siendo superior en el caso de la presencia.

Para la variable *eficiencia* los resultados son reveladores. Aunque en ningún caso la existencia del mecanismo de usabilidad perjudica la eficiencia de los usuarios no siempre se observa una mejora determinante en el test ANOVA. En el caso del mecanismo *Abortar Operación* sí se observa una mejora significativa en la eficiencia del usuario (tanto en su rapidez como en el nivel de interacción con la aplicación) en la realización de la tarea cuando este mecanismo está presente. En el caso del mecanismo *Retroalimentación de Progreso* la eficiencia sólo es significativa cuando se mide como el nivel de interacción del usuario con la aplicación. Sin embargo, para el mecanismo *Preferencias* se observa que el orden de realización de las tareas es un factor decisivo.

De esta investigación se desprende que es verdaderamente importante tener en cuenta, desde el diseño, los mecanismos de usabilidad de *Abortar Operación*, *Retroalimentación de Progreso* y *Preferencias* para conseguir una mejora significativa en la satisfacción del usuario interaccionando con el sistema. También es importante disponer de estos mecanismos para obtener una mejora en la eficiencia de los usuarios.

Palabras clave: *usabilidad, estudio empírico, eficiencia, satisfacción, experimento verdadero, Ingeniería del Software.*

Abstract

Usability is an increasingly important part of every kind of software application, and especially of highly interactive web services. However, there are no experimental studies offering satisfactory statistical evidence regarding what impact the implementation of different usability mechanisms really has on application end user efficiency and satisfaction.

This research was designed to specifically address this question. To be precise, an empirical study was conducted in order to determine how the adoption of usability mechanisms affects efficiency and user satisfaction. For this purpose, three usability mechanisms -*Abort Operation*, *Progress Feedback* and *Preferences*- were evaluated on an online shop, whose main purpose is the management of online commerce.

This study is implemented using an experiment that has a *between subject* design with a total of three factors. The first is the adoption/non-adoption of the mechanism under study. The second is the group. There are four groups {A, B, C, D}. Only one of the mechanisms is available in each of groups A, B and C, whereas all the mechanisms are adopted in D. In other words, the mechanism will be enabled in half of the cases and disabled in the other half. The third factor is the order in which the user performs the different tasks (a permutation of their order).

A number of sessions with 100 experimental subjects were held according to the experimental design working on the constructed application. Each experimental subject was asked to perform the three experimental tasks, one task per usability mechanism. The time that subjects take to perform each task and the number of clicks that they use is measured automatically as they complete the experiment. These two metrics measure how efficient participants are at performing each task. Besides, they are asked to complete a questionnaire at the end of each task. This is used to measure user satisfaction.

At the end of the experiment, a carefully designed statistical analysis is conducted. Basic descriptive statistics (average, median, typical deviation, skewness and kurtosis, etc.) and boxplots for every permutation provide a simplistic representation of the measures and an overview of the information. This analysis identified some unfeasible values –data that are not outliers but corroborate that some subjects failed to complete the activity successfully–, which are removed from the original sample.

Then, after testing for the normality of the residuals and the homogeneity of the variances, it is feasible to perform an ANOVA test on this filtered sample for both response variables (efficiency and satisfaction) and each usability mechanism (*Abort Operation*, *Progress Feedback* and *Preferences*).

In the case of the *satisfaction* response variable, the results are conclusive: there is a clear difference in all three cases between the adoption or non-adoption of the usability mechanism with respect to the final satisfaction of subjects, which is greater in the case of adoption.

The results for the efficiency variable are revealing. Even though the adoption of the usability mechanism never decreases the efficiency of users, the ANOVA test does not always identify a definite improvement. In the case of the *Abort Operation* mechanism, there is found to be a significant improvement in the user *efficiency* (with respect to

both speed and level of interaction with the application) when this mechanism is adopted to perform the task. On the other hand, *Progress Feedback* only causes a significant improvement when it is measured as the level of interaction of the user with the application. However, the order in which the tasks are scheduled is a critical factor for the *Preferences* mechanism.

We can conclude from this research that it is extremely important to take into account different types of usability mechanisms as of design time. *Abort Operation*, *Progress Feedback* and *Preferences* can ensure a significant increase in user satisfaction when interacting with the system. Note that the adoption of these features also has a positive impact on efficiency.

Keywords: *usability, empirical study, efficiency, satisfaction, true experiment, Software Engineering.*

Índice

| | |
|---|-----------|
| CAPÍTULO 1. INTRODUCCIÓN..... | 1 |
| 1.1. Visión General de la Investigación | 1 |
| 1.2. Usabilidad en la Ingeniería del Software | 2 |
| 1.3. Problema de Investigación..... | 4 |
| 1.4. Método de la Investigación..... | 5 |
| 1.5. Estructura del Trabajo..... | 6 |
| 1.6. Contribución y Publicación Derivada | 7 |
| CAPÍTULO 2. ESTADO DE LA CUESTIÓN | 8 |
| 2.1. Método de Investigación | 8 |
| 2.2. Preguntas de Investigación | 9 |
| 2.3. Proceso de Búsqueda y Selección | 9 |
| 2.4. Criterios de Inclusión-Exclusión | 10 |
| 2.5. Resultados..... | 12 |
| CAPÍTULO 3. EXPERIMENTO VERDADERO | 15 |
| 3.1. Introducción..... | 15 |
| 3.2. Diseño Experimental | 15 |
| 3.3. Objetivos de Investigación | 17 |
| 3.4. Hipótesis de Investigación | 17 |
| 3.5. Variables Respuesta..... | 18 |
| 3.6. Sujetos | 18 |
| 3.7. Instrumentación..... | 19 |
| 3.8. Procedimiento de Recolección de Datos | 21 |
| 3.8.1. Métricas | 21 |
| 3.8.2. Cuestionarios..... | 22 |
| 3.9. Validez Interna y Externa | 23 |
| 3.9.1. Validez Interna | 23 |
| 3.9.2. Validez Externa | 24 |
| 3.10. Análisis de Datos | 24 |
| 3.11. Análisis Descriptivo | 26 |
| 3.11.1. Eficiencia | 26 |
| 3.11.1.1. Preferencias | 26 |
| 3.11.1.2. Retroalimentación de Progreso | 29 |
| 3.11.1.3. Abortar Operación | 31 |
| 3.11.2. Satisfacción..... | 33 |
| 3.11.2.1. Preferencias | 33 |
| 3.11.2.2. Retroalimentación de Progreso | 34 |
| 3.11.2.3. Abortar Operación | 35 |
| 3.12. ANOVA | 36 |
| 3.12.1. Eficiencia | 36 |
| 3.12.1.1. Preferencias | 36 |
| 3.12.1.2. Retroalimentación de Progreso | 38 |
| 3.12.1.3. Abortar Operación | 39 |
| 3.12.2. Satisfacción..... | 41 |
| 3.12.2.1. Preferencias | 41 |
| 3.12.2.2. Retroalimentación de Progreso | 41 |
| 3.12.2.3. Abortar Operación | 42 |
| CAPÍTULO 4. DISCUSIÓN DE RESULTADOS..... | 44 |
| 4.1. Discusión | 44 |

| | |
|---|-----------|
| 4.2. Interpretación | 45 |
| CAPÍTULO 5. CONCLUSIONES Y TRABAJOS FUTUROS..... | 48 |
| 5.1. Conclusiones | 48 |
| 5.2. Limitaciones del Trabajo | 49 |
| 5.3. Trabajos Futuros | 49 |
| REFERENCIAS..... | 50 |
| GLOSARIO | 53 |
| ANEXOS..... | 54 |
| Anexo A. Evolución del Diseño..... | 54 |
| A.1. Diseño Full Factorial | 54 |
| A.2. Primera Versión del Diseño Ortogonal | 56 |
| A.3. Segunda Versión del Diseño Ortogonal | 57 |
| A.4. Diseño Between Subjects del Experimento Piloto | 58 |
| Anexo B. Experimento Piloto | 61 |
| B.1. Introducción..... | 61 |
| B.2. Análisis de Datos..... | 61 |
| B.3. Análisis Descriptivo | 62 |
| B.3.1. Eficiencia | 62 |
| B.3.2. Satisfacción | 64 |
| B.4. ANOVA | 65 |
| B.4.1. Eficiencia | 65 |
| B.4.1.1. Número de Clicks | 65 |
| B.4.1.2. Tiempo | 66 |
| B.4.2. Satisfacción | 66 |
| B.5. Discusión e Interpretación de Resultados | 67 |
| Anexo C. Diseño e Implementación de las Tareas | 68 |
| C.1. Aplicación <i>QuickStore</i> | 68 |
| C.1.1. Descripción de la Aplicación..... | 68 |
| C.1.2. Arquitectura | 68 |
| C.2. Mecanismo de Usabilidad Abortar Operación..... | 69 |
| C.3. Mecanismo de Usabilidad Retroalimentación de Progreso..... | 70 |
| C.4. Mecanismo de Usabilidad Preferencias..... | 72 |
| Anexo D. Diagramas de Caja del Experimento Verdadero | 75 |
| D.1. Eficiencia..... | 75 |
| D.1.1. Número de Clicks | 75 |
| D.1.1.1. Preferencias..... | 75 |
| D.1.1.2. Retroalimentación de Progreso..... | 76 |
| D.1.1.3. Abortar Operación..... | 78 |
| D.1.2. Tiempo | 79 |
| D.1.2.1. Preferencias..... | 79 |
| D.1.2.2. Retroalimentación de Progreso..... | 81 |
| D.1.2.3. Abortar Operación..... | 82 |
| D.2. Satisfacción..... | 84 |
| D.2.1. Preferencias | 84 |
| D.2.2. Retroalimentación de Progreso | 85 |
| D.2.3. Abortar Operación | 87 |
| Anexo E. Test de Tukey..... | 89 |
| Anexo F. Cuestionario de Familiaridad..... | 92 |

Índice de Tablas

| | |
|--|----|
| Tabla 1: Resumen de características y mecanismos de usabilidad | 3 |
| Tabla 2: Elección de mecanismos de usabilidad | 4 |
| Tabla 3: Estudios encontrados en los cinco motores de búsqueda | 10 |
| Tabla 4: Detalle del filtrado realizado desde los estudios encontrados hasta obtener los estudios primarios | 12 |
| Tabla 5: Matriz de grupos | 16 |
| Tabla 6: Matriz con el orden de exposición de cada mecanismo de usabilidad | 16 |
| Tabla 7: Matriz de asignación de agrupaciones con órdenes de exposición..... | 16 |
| Tabla 8: Hipótesis de investigación | 18 |
| Tabla 9: Información referente a compras..... | 19 |
| Tabla 10: Columnas con los datos de las métricas observadas..... | 21 |
| Tabla 11: Número de filas en las tablas de métricas..... | 21 |
| Tabla 12: Columnas de las tablas con datos de los cuestionarios..... | 22 |
| Tabla 13: Número de filas de las tablas con información de los cuestionarios..... | 23 |
| Tabla 14: Elementos considerados en los test ANOVA | 26 |
| Tabla 15: Estadísticos descriptivos para el mecanismo Preferencias en el estudio de Eficiencia | 27 |
| Tabla 16: Estadísticos descriptivos para el mecanismo Retroalimentación de Progreso en el estudio de Eficiencia..... | 29 |
| Tabla 17: Estadísticos descriptivos para el mecanismo Abortar Operación en el estudio de Eficiencia | 31 |
| Tabla 18: Estadísticos descriptivos para el mecanismo Preferencias para el estudio de Satisfacción..... | 33 |
| Tabla 19: Estadísticos descriptivos para el mecanismo Retroalimentación de Progreso para el estudio de Satisfacción..... | 34 |
| Tabla 20: Estadísticos descriptivos para el mecanismo Abortar Operación para el estudio de Satisfacción..... | 35 |
| Tabla 21: Distintos niveles de significación considerados..... | 36 |
| Tabla 22: Test de normalidad y homocedasticidad variable Tiempo MU-PRE..... | 37 |
| Tabla 23: Test de ANOVA para la variable Tiempo en MU-PRE | 37 |
| Tabla 24: Test de normalidad y homocedasticidad variable Nº de Clicks MU-PRE | 37 |
| Tabla 25: Test de ANOVA para la variable Nº de Clicks en MU-PRE..... | 38 |
| Tabla 26: Test de normalidad y homocedasticidad variable Tiempo MU-PRO | 39 |
| Tabla 27: Test de ANOVA para la variable Tiempo en MU-PRO..... | 39 |
| Tabla 28: Test de normalidad y homocedasticidad variable Nº de Clicks MU-PRO | 39 |
| Tabla 29: Test de ANOVA para la variable Nº de Clicks en MU-PRO..... | 39 |
| Tabla 30: Test de normalidad y homocedasticidad variable Tiempo MU-ABO..... | 40 |
| Tabla 31: Test de ANOVA para la variable Tiempo en MU-ABO | 40 |
| Tabla 32: Test de normalidad y homocedasticidad variable Nº de Clicks MU-ABO | 40 |
| Tabla 33: Test de ANOVA para la variable Nº de Clicks en MU-ABO..... | 40 |
| Tabla 34: Test de normalidad y homocedasticidad para la Satisfacción en MU-PRE..... | 41 |
| Tabla 35: Test de ANOVA para la Satisfacción en MU-PRE | 41 |
| Tabla 36: Test de normalidad y homocedasticidad para la Satisfacción en MU-PRO | 42 |
| Tabla 37: Test de ANOVA para la Satisfacción en MU-PRO..... | 42 |
| Tabla 38: Test de normalidad y homocedasticidad para la Satisfacción en MU-ABO | 42 |
| Tabla 39: Test de ANOVA para la Satisfacción en MU-ABO | 43 |
| Tabla 40: Resumen de resultados del experimento con ANOVA | 44 |

| | |
|---|----|
| Tabla 41: Matriz de grupos Full Factorial | 55 |
| Tabla 42: Matriz con el orden de exposición de cada MU del diseño Full Factorial | 55 |
| Tabla 43: Matriz de asignación de agrupaciones con órdenes de exposición del diseño Full Factorial..... | 55 |
| Tabla 44: Matriz de grupos versión 1 del Diseño Ortogonal | 56 |
| Tabla 45: Matriz con el orden de exposición por Cuadrados Latinos | 56 |
| Tabla 46: Matriz de asignación de agrupaciones con órdenes de exposición..... | 56 |
| Tabla 47: Órdenes de exposición válidos de la segunda versión del Diseño Ortogonal | 57 |
| Tabla 48: Matriz de asignación de agrupaciones con órdenes de exposición..... | 57 |
| Tabla 49: Matriz de grupos Between Subject | 59 |
| Tabla 50: Matriz con el orden de exposición con secuencias lógicas..... | 59 |
| Tabla 51: Matriz de asignación de agrupaciones con órdenes de exposición..... | 59 |
| Tabla 52: Elementos considerados en los test ANOVA para el experimento piloto | 62 |
| Tabla 53: Estadísticos descriptivos para la Eficiencia en el experimento piloto | 63 |
| Tabla 54: Estadísticos descriptivos para la Satisfacción en el experimento piloto | 64 |
| Tabla 55: Resultados Eficiencia - Nº de Clicks del experimento piloto | 66 |
| Tabla 56: Resultados Eficiencia - Tiempo del experimento piloto | 66 |
| Tabla 57: Resultados Satisfacción del experimento piloto..... | 67 |
| Tabla 58: Contexto, tareas y secuencia de acciones del MU-ABO | 69 |
| Tabla 59: Diseño del cuestionario de satisfacción post-tarea del MU-ABO | 70 |
| Tabla 60: Contexto, tareas y secuencia de acciones del MU-PRO | 71 |
| Tabla 61: Diseño del cuestionario de satisfacción post-tarea del MU-PRO | 72 |
| Tabla 62: Contexto, tareas y secuencia de acciones del MU-PRE | 73 |
| Tabla 63: Diseño del cuestionario de satisfacción post-tarea del MU-PRE..... | 74 |

Índice de Figuras

| | |
|--|----|
| Figura 1: Fases de la investigación | 5 |
| Figura 2: Proceso Systematic Mapping Study | 9 |
| Figura 3: Método estadístico | 24 |
| Figura 4: Diagrama de cajas para el Número de Clicks en el mecanismo Preferencias | 28 |
| Figura 5: Diagrama de cajas para el Tiempo, en segundos, en el mecanismo Preferencias | 28 |
| Figura 6: Diagrama de cajas para el Número de Clicks en el mecanismo Retroalimentación de Progreso | 30 |
| Figura 7: Diagrama de cajas para el Tiempo en el mecanismo Retroalimentación de Progreso | 30 |
| Figura 8: Diagrama de cajas para el Número de Clicks en el mecanismo Abortar Operación | 32 |
| Figura 9: Diagrama de cajas para el Tiempo en el mecanismo Abortar Operación | 32 |
| Figura 10: Diagrama de cajas para la Satisfacción en el mecanismo Preferencias | 33 |
| Figura 11: Diagrama de cajas para la Satisfacción en el mecanismo Retroalimentación de Progreso | 34 |
| Figura 12: Diagrama de cajas para la Satisfacción en el mecanismo Abortar Operación | 35 |
| Figura 13: Diagrama de cajas, datos agrupados, para la variable N° de Clicks en el experimento piloto..... | 63 |
| Figura 14: Diagrama de cajas, datos agrupados, para la variable Tiempo en el experimento piloto..... | 64 |
| Figura 15: Diagrama de cajas, datos agrupados, para la variable Satisfacción en el experimento piloto..... | 65 |
| Figura 16: Arquitectura MVC de Java | 68 |
| Figura 17: Árbol de escenarios de aplicación del MU-ABO | 69 |
| Figura 18: Interfaz de tarea para el MU-ABO | 70 |
| Figura 19: Árbol de escenarios de aplicación del MU-PRO | 71 |
| Figura 20: Interfaz de tarea para el MU-PRO | 72 |
| Figura 21: Interfaz de tarea para el MU-PRE..... | 73 |
| Figura 22: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRE y N° de Clicks | 75 |
| Figura 23: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRE y N° de Clicks | 76 |
| Figura 24: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRO y N° de Clicks..... | 77 |
| Figura 25: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRO y N° de Clicks..... | 77 |
| Figura 26: Diagrama de cajas Presencia/Ausencia y Grupo para MU-ABO y N° de Clicks | 78 |
| Figura 27: Diagrama de cajas Presencia/Ausencia y Orden para MU-ABO y N° de Clicks | 79 |
| Figura 28: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRE y Tiempo | 80 |
| Figura 29: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRE y Tiempo | 80 |
| Figura 30: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRO y Tiempo | 81 |
| Figura 31: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRO y Tiempo | 82 |
| Figura 32: Diagrama de cajas Presencia/Ausencia y Grupo para MU-ABO y Tiempo | 83 |
| Figura 33: Diagrama de cajas Presencia/Ausencia y Orden para MU-ABO y Tiempo | 83 |
| Figura 34: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRE y Satisfacción | 84 |
| Figura 35: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRE y Satisfacción | 85 |
| Figura 36: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRO y Satisfacción..... | 86 |
| Figura 37: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRO y Satisfacción..... | 86 |
| Figura 38: Diagrama de cajas Presencia/Ausencia y Grupo para MU-ABO y Satisfacción | 87 |
| Figura 39: Diagrama de cajas Presencia/Ausencia y Orden para MU-ABO y Satisfacción | 88 |
| Figura 40: Test de Tukey para el factor Orden y la variable Tiempo MU-PRE..... | 89 |
| Figura 41: Test de Tukey para el factor Orden y la variable N° de Clicks MU-PRE..... | 90 |
| Figura 42: Test de Tukey para el factor Orden y la variable Tiempo MU-PRO..... | 90 |

| | |
|---|----|
| Figura 43: Test de Tukey para el factor Orden y la variable Nº de Clicks MU-PRO | 91 |
| Figura 44: Test de Tukey para el factor Orden y la variable Tiempo MU-ABO..... | 91 |
| Figura 45: Cuestionario de Familiaridad | 92 |

CAPÍTULO 1. INTRODUCCIÓN

1.1. Visión General de la Investigación

La usabilidad es un atributo de calidad que, según el estándar internacional [ISO/IEC 25010, 2010], se define como la capacidad del producto software de ser entendido, aprendido, usado y atractivo para el usuario cuando se usa bajo condiciones específicas. Según [Juristo et al., 2007a; Winter & Rönkkö, 2010; Battey, 2001; Donahue, 2001; Black, 2002] tener en cuenta este atributo de calidad proporciona varios beneficios añadidos: mejoras en la relación coste-beneficio con su consiguiente aumento en los ingresos de proyectos software, mejoras en la productividad del equipo de trabajo, facilidad en el mantenimiento de los productos, etc.

En la literatura se encuentran recomendaciones para mejorar la usabilidad de los sistemas software [Bass et al., 2001]. Sin embargo, no existen datos empíricos que demuestren cómo estas recomendaciones contribuyen positiva o negativamente en cada atributo de usabilidad. En este contexto, el objetivo de este trabajo es la obtención de evidencias empíricas sobre el impacto de la inclusión de mecanismos de usabilidad en un sistema software, más concretamente en los atributos de usabilidad eficiencia y satisfacción.

Esta investigación se centra en considerar tres mecanismos de usabilidad en un entorno web con grandes características interactivas. Concretamente, los mecanismos de Preferencias (PRE), Abortar Operación (ABO) y Retroalimentación de Progreso (PRO). Debido a las características interactivas del sistema software que se utilizará, dichos mecanismos de calidad pueden llegar a ser críticos por su gran impacto en el diseño del software tal y como se destaca en [Juristo et al., 2007b].

Para realizar el análisis empírico se utiliza un experimento verdadero. Estos experimentos, que también se conocen como experimentos controlados, manipulan las variables independientes consideradas para ver sus efectos sobre las variables dependientes en una situación de control. En el tercer capítulo de este trabajo, se detalla, de forma pormenorizada, el diseño del experimento y las razones que han llevado a determinar que se realizará dicho experimento verdadero.

La obtención de tales evidencias es fundamental, ya que incorporar estos mecanismos en un sistema software requiere de un considerable esfuerzo de diseño materializado en la creación de componentes específicos (por ejemplo, clases, métodos, etc.) que han de incorporarse en los modelos de desarrollo funcionales. Por lo tanto, la obtención de estas evidencias justificará o no este esfuerzo adicional requerido. Más aún, resulta necesaria la realización de una investigación que determine los beneficios que puedan aportar la inclusión de los mecanismos de usabilidad en sistemas software web.

1.2. Usabilidad en la Ingeniería del Software

Hoy en día, la usabilidad es uno de los atributos clave en el desarrollo de software [Ferré et al., 2001]. El estándar internacional [ISO 9241-11, 1998] define este concepto como el grado en que un producto puede ser usado por determinados usuarios para conseguir objetivos específicos con eficacia, eficiencia y satisfacción en un contexto de uso.

Desde la Ingeniería del Software (IS), en un principio, se consideraba que la usabilidad era un atributo que estaba únicamente relacionado con la interfaz de usuario (IU) y, por tanto, solamente repercutía en los requisitos no funcionales. Por esta razón, se abordaba utilizando estrategias de diseño como el Modelo Vista Controlador (MVC) o el modelo Presentación Abstracción Control (PAC). Ambos diseños separan la parte de presentación gráfica de la funcionalidad propia del sistema.

Sin embargo, esta opinión está cambiando debido a las conclusiones de distintos estudios. Por ejemplo, en [Bass & John, 2003] se presenta un conjunto de escenarios de usabilidad, en los que se evidencia que la estrategia de separación de la IU no es suficiente para obtener un sistema usable. En [Juristo et al., 2007a] se obtuvieron, de forma empírica, evidencias que confirmaban la relación entre la usabilidad y el diseño del software. Aquí, se determinó también que el costo de repetir el trabajo para obtener un nivel aceptable de usabilidad sería mucho más alto del esperado de acuerdo a la hipótesis de separación entre la IU y el *core* del sistema software ya que, como estamos comentado, pueden existir atributos de usabilidad que afecten directamente a la funcionalidad del sistema.

Es importante tener en cuenta que, dependiendo de los autores, se realizan distintas recomendaciones sobre la usabilidad en un sistema software con impacto sobre el diseño. En este trabajo se tendrá en cuenta la clasificación realizada en [Juristo et al., 2007b] donde se presentan distintas características funcionales de usabilidad (CFU). Las CFU que obtienen una clasificación de impacto alto en el diseño del software, en este estudio son las siguientes:

- **Adaptación** al perfil de usuario: permitir al usuario adaptar el sistema a sus propias características.
- **Agregación de comandos**: expresar posibles acciones a ser tomadas con el software a través de comandos que puedan ser construidos desde pequeñas partes.
- **Asistentes**: ayudar al usuario a hacer tareas que involucran su interacción con el sistema y que requieren más de un paso.
- **Ayudas**: proveer al usuario con ayudas para completar tareas dentro de la aplicación.
- **Cancelar**: permitir al usuario cancelar acciones que están en curso a varios niveles.
- **Deshacer**: permitir al usuario deshacer acciones a varios niveles.
- **Prevención y corrección de errores introducidos por el usuario**: se refiere a validaciones de entrada y corrección pronta de errores.
- **Retroalimentación**: mantener informado al usuario sobre lo que sucede en el sistema.

En el mismo trabajo se realiza una subcategorización de las características anteriores. Dicha clasificación se detalla en la Tabla 1.

| Característica de Usabilidad | Mecanismos de Usabilidad | Objetivo |
|--|-------------------------------------|---|
| Agregación de Comandos | Agregación de comandos | Expresar posibles acciones a ser tomadas con el software a través de comandos que puedan ser contruidos desde pequeñas partes. |
| Asistentes | Ejecución paso a paso | Ayudar a los usuarios a hacer tareas que requieren diferentes pasos con entradas del usuario y corregir tal información. |
| Ayudas | Ayudas multinivel | Proveer diferentes niveles de ayuda para diferentes usuarios. |
| Deshacer/Cancelar | Abortar operación | Cancelar la ejecución de una acción o de toda la aplicación. |
| | Deshacer global | Deshacer acciones del sistema a varios niveles. |
| | Deshacer sobre un objeto específico | Deshacer varias acciones sobre un objeto. |
| | Regresar | Regresar a un estado particular en una secuencia de ejecución de comandos. |
| Perfil de Usuario | Espacio personalizado de objetos | Grabar cada una de las opciones del usuario al usar la interfaz del sistema. |
| | Favoritos | Guardar ciertos lugares de interés para el usuario. |
| | Preferencias | Grabar cada una de las opciones del usuario para el uso de las funciones del sistema. |
| Prevención/Corrección de Errores Introducidos por el Usuario | Entrada estructurada de texto | Ayuda para prevenir que el usuario introduzca errores al ingresar datos. |
| Retroalimentación | Alertas | Informar al usuario de cualquier acción con consecuencias importantes. |
| | Estado del sistema | Informar al usuario acerca del estado interno del sistema. |
| | Interacción | Informar al usuario que el sistema ha registrado una interacción del usuario, por ejemplo que ha pasado el ratón encima de alguna acción. |
| | Retroalimentación de progreso | Informar al usuario que el sistema está procesando una acción que tomará algún tiempo para completarse. |

Tabla 1: Resumen de características y mecanismos de usabilidad

Para la realización de este trabajo, no se pretende tener en cuenta todas las características de usabilidad presentadas. Si así fuese, surgirían problemas importantes a la hora de identificar las variables experimentales y de construir los escenarios del experimento que se llevará a cabo.

Es por ello que existe la necesidad de determinar cuáles son las características más importantes para su estudio. En los trabajos de [Juristo et al., 2007a; Juristo et al., 2007b] se consideran que los tres mecanismos de usabilidad de mayor impacto son:

- Abortar Operación
- Preferencias
- Retroalimentación de Progreso.

Los criterios que se utilizaron en los trabajos citados para llegar a esta conclusión han sido: nivel de impacto sobre el diseño en cuanto a número de funcionalidades que serían afectadas en las aplicaciones a desarrollar, facilidad de reconocimiento por parte de un usuario cuando use el sistema y facilidad de evaluación desde el punto de vista de las recomendaciones de la disciplina de la Interacción Persona-Ordenador.

La Tabla 2 presenta la evaluación obtenida para los distintos mecanismos de usabilidad según los criterios mencionados. Los mecanismos destacados en color verde (Tabla 1 y Tabla 2) son los que estarán presentes en este trabajo: Abortar Operación, Retroalimentación de Progreso y Preferencias. Estos tres mecanismos tienen un impacto alto y, además, son fáciles de reconocer y evaluar (Tabla 2).

| Mecanismo de Usabilidad | Criterio | | |
|-------------------------------------|------------------|-----------------------------|-------------------------|
| | Nivel de Impacto | Facilidad de Reconocimiento | Facilidad de Evaluación |
| Estado del sistema | Bajo | | |
| Interacción | Medio | X | X |
| Alertas | Bajo | | |
| Retroalimentación de progreso | Alto | X | X |
| Deshacer global | Bajo | X | X |
| Deshacer sobre un objeto específico | Bajo | X | X |
| Abortar operación | Alto | X | X |
| Regresar | Bajo | X | X |
| Entrada estructurada de texto | Alto | X | |
| Ejecución paso a paso | Bajo | X | X |
| Preferencias | Alto | X | X |
| Espacio personalizado de objetos | Bajo | X | X |
| Favoritos | Bajo | X | X |
| Ayudas multinivel | Bajo | X | X |
| Agregación de comandos | Bajo | X | |

Tabla 2: Elección de mecanismos de usabilidad

1.3. Problema de Investigación

Como se ha presentado anteriormente, esta investigación se centrará en medir la eficiencia y la satisfacción de los usuarios considerando la presencia o ausencia de los mecanismos de usabilidad Retroalimentación de Progreso, Abortar Operación y Preferencias sobre un entorno web con altas características interactivas.

Para comprobar el impacto de la presencia o ausencia de los mecanismos citados en la eficiencia y satisfacción, se usará una muestra de 100 usuarios que deberán realizar tres tareas sobre una aplicación web diseñada específicamente para este fin. Concretamente se utilizará la aplicación *QuickStore* que es una herramienta de compra online de distintos tipos de artículos.

En este estudio empírico se evaluará la eficiencia y la satisfacción de los sujetos al enfrentarse a las tareas que conforman el test de usabilidad. En el Anexo C, se puede encontrar la descripción de la aplicación web y, además, se detalla la propuesta de interfaz gráfica para guiar el experimento.

Los objetivos generales de esta investigación son:

- ✓ Estudiar en profundidad la literatura y los trabajos existentes sobre cómo realizar evaluaciones empíricas para medir atributos de usabilidad y determinar el estado del arte de las investigaciones.
- ✓ Formalizar el análisis del estudio para obtener las evidencias necesarias para construir la investigación y diseñar el experimento.
- ✓ Estudiar empíricamente la influencia de la presencia o la ausencia del mecanismo *Retroalimentación de Progreso* en la usabilidad de la aplicación.
- ✓ Estudiar empíricamente la influencia de la presencia o la ausencia del mecanismo *Abortar Operación* en la usabilidad de la aplicación.
- ✓ Estudiar empíricamente la influencia de la presencia o la ausencia del mecanismo *Preferencias* en la usabilidad de la aplicación.

Para conseguir los tres últimos objetivos, se tendrán en cuenta dos variables respuesta en el análisis: la eficiencia y la satisfacción, mediante la operacionalización de la variable usabilidad. En la sección 3.5 se definen estas variables y en las secciones 3.3 y 3.4 se presentan los objetivos concretos y las hipótesis de investigación que se tendrán en cuenta.

1.4. Método de la Investigación

Para lograr el objetivo planteado en esta investigación, se propone un método que se realizará en tres fases, ver Figura 1. En la primera fase, se diseñará el experimento verdadero para comprobar el impacto de los mecanismos de usabilidad seleccionados: Abortar Operación, Retroalimentación de Progreso y Preferencias, en la eficiencia y la satisfacción de los usuarios utilizando la aplicación web *QuickStore*, diseñada para este trabajo (ver anexos A y C). En la segunda fase, se llevará a cabo el experimento con los 100 sujetos. Por último, y en la tercera fase, se analizarán los resultados obtenidos en el experimento para poder realizar las conclusiones que sean pertinentes.

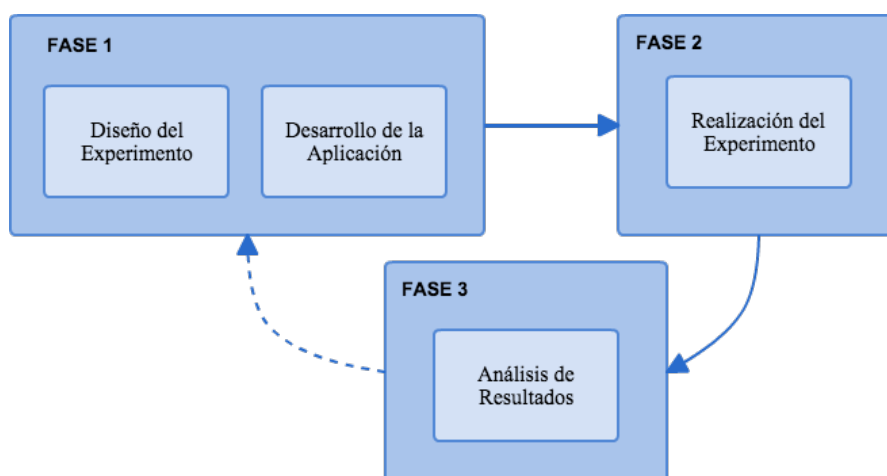


Figura 1: Fases de la investigación

Cabe destacar que estas tres fases se realizarán de forma iterada pues, en primer lugar, se diseñará un experimento piloto (ver anexo B) para comprobar si son necesarias modificaciones antes de presentar el experimento verdadero a los 100 sujetos implicados.

1.5. Estructura del Trabajo

Este trabajo se divide en cinco capítulos, los apartados de Referencias y Glosarios y cuenta además con seis Anexos. A continuación, se describe el contenido de cada uno de ellos.

El primer capítulo sirve de presentación de la investigación que se pretende llevar a cabo. En él se realiza una visión general de la investigación y se plantea el problema a investigar y los objetivos del trabajo, así como también el método que se seguirá para aproximarse a la solución.

En el segundo capítulo se aborda el estado de la cuestión. Para obtener toda la información ya existente sobre el tema de investigación se ha realizado un *Systematic Mapping Study* (SMS) con el fin de realizar el proceso de estudio de la literatura de forma metódica y eficaz.

En el tercer capítulo se detalla, de forma pormenorizada, el experimento verdadero llevado a cabo. Aquí, por una parte, se presentan los objetivos y las hipótesis del experimento. Además, se describen las variables respuesta que se medirán. Por otra parte, también se analizan los sujetos que han intervenido en el experimento verdadero, los instrumentos utilizados, los datos recolectados, la validez interna y externa del estudio llevado a cabo y, por último, el análisis de datos y los resultados obtenidos en la evaluación.

En el cuarto capítulo se discuten los resultados obtenidos y su interpretación.

En el quinto capítulo se presentan las conclusiones y el trabajo futuro.

Seguidamente se presenta un apartado con las referencias utilizadas durante todo el proceso de investigación y un glosario con definiciones estadísticas aclaratorias.

En el anexo A se describe la evolución del diseño del experimento desde su fase inicial hasta la versión utilizada en el experimento verdadero final.

En el anexo B se detalla el experimento piloto llevado a cabo. Se realiza un análisis similar al realizado en el tercer capítulo para el experimento verdadero.

En el anexo C se detalla el diseño y la implementación de las tareas del experimento verdadero final.

En el anexo D se comentan los diagramas de caja del experimento verdadero que complementan a los recogidos en el tercer capítulo.

En el anexo E aparecen los diagramas que recogen los resultados de los test de Tukey que complementan los test ANOVA cuando el mecanismo orden es determinante.

En el anexo F se presenta el cuestionario de familiaridad utilizado tanto en el experimento piloto como en el experimento verdadero final.

1.6. Contribución y Publicación Derivada

El presente trabajo de investigación ha dado lugar a la siguiente publicación derivada en congreso internacional:

Herea Bouza Leirós, Silvia T. Acuña. (2016). How does the Abort Operation, Progress Feedback and Preferences Usability Mechanisms Influence User Satisfaction and Efficiency? A Controlled Experiment. Live Study Contribution, Research Method Track, 22nd International Working Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2016). Göteborg, Sweden, March 14-17, pp. 1-4.

El experimento diseñado en este trabajo se ejecutará durante el REFSQ 2016 por parte de los aproximadamente 80-100 participantes del evento y las autoras presentarán, de modo sintético, el diseño de investigación y los resultados de la réplica del experimento de este trabajo en la sesión plenaria de la conferencia.

CAPÍTULO 2. ESTADO DE LA CUESTIÓN

En el presente capítulo se abordan las propuestas existentes en la literatura sobre la realización de distintos experimentos empíricos para medir atributos de usabilidad. Con el objetivo de realizar una comparación entre estos experimentos se utilizará el método *Systematic Mapping Study* (SMS) presentado en la sección 2.1, para realizar esta búsqueda de forma metódica y ágil. Las secciones 2.2 y 2.3 se refieren a las preguntas de investigación, a las cadenas y los motores de búsqueda utilizados en el SMS, respectivamente. En la sección 2.4 se definen los criterios de inclusión-exclusión necesarios para la correcta realización del SMS. Finalmente, en la sección 2.5 se presentan los resultados del SMS para determinar el estado de la cuestión.

2.1. Método de Investigación

Una Revisión Sistemática de la Literatura (en inglés *Systematic Literature Review*, SLR) es un medio para identificar, evaluar e interpretar toda la investigación disponible sobre una o varias preguntas de investigación o sobre algún área de interés [Kitchenham, 2007].

Existen distintas razones para llevar a cabo una SLR. Entre ellas destacan las siguientes: resumir las evidencias existentes respecto a una tecnología o tema; identificar errores, vacíos o temas incompletos en los estudios concluidos para encontrar áreas que potencialmente podrían ser investigadas en el futuro; o proveer de una base apropiada para el inicio de una nueva investigación.

Si bien una SLR tiene un método completamente definido para ser llevada a cabo, es un tipo de estudio no recomendable si el tema del que se pretende encontrar información es complejo y extenso. Para estos casos, es más conveniente realizar un SMS [Petersen et al., 2007]. Estos estudios también se conocen como *Scoping Studies*.

Los SMS están diseñados para proporcionar una visión general de un área de investigación. Permiten, como las SLR, la identificación de núcleos de investigación desiertos para focalizar las futuras investigaciones. Son más útiles que las SLR por realizarse mediante un proceso más dinámico y ágil que el seguido en las primeras. En los trabajos de [Kitchenham, 2004; Acuña & Castro, 2012] se pueden encontrar distintos ejemplos.

La Figura 2 representa el proceso que sigue un SMS. Concretamente, se muestra el proceso utilizado en este trabajo para la obtención de los estudios necesarios para el inicio de la investigación que se realiza aquí.

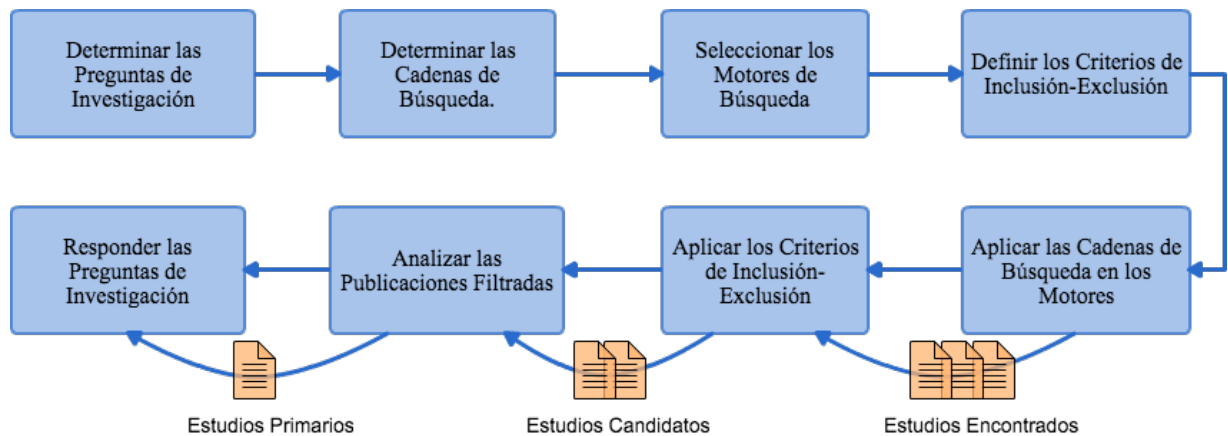


Figura 2: Proceso Systematic Mapping Study

2.2. Preguntas de Investigación

Tal y como muestra la Figura 2, para seguir el método SMS el primer paso que se debe realizar es determinar cuáles serán las preguntas de investigación. Una pregunta de investigación será una cuestión que se desea resolver antes de empezar un estudio, cuya respuesta ayuda a formar las bases de partida para el mismo.

Teniendo en cuenta los objetivos del trabajo se han determinado las siguientes cuestiones:

- [Q1] ¿Cuáles son los temas, relacionados con la usabilidad, objeto de investigación empírica?
- [Q2] ¿Qué tipos de estudios cuantitativos se realizan en usabilidad?

2.3. Proceso de Búsqueda y Selección

Una vez se han identificado las preguntas de investigación, necesitamos toda la información que haga falta para poder responderlas. Para ello, debemos reescribir las mismas de una forma concisa con el fin de poder utilizar los distintos motores de búsqueda.

Estos motores de búsqueda aceptan una o varias cadenas de búsqueda que contengan las distintas combinaciones de las palabras que se espera que aparezcan en la publicación destino. Actualmente, los motores de búsqueda que se utilizarán en este SMS proporcionan una búsqueda avanzada que permite la utilización de distintos operadores lógicos. Concretamente, se utilizarán los operadores “AND” y “OR” a la hora de crear las distintas cadenas de búsqueda. El uso de estos dos operadores se sugiere en [Kitchenham, 2004].

Una segunda cuestión a resolver es dónde se utilizarán las cadenas de búsqueda. Éstas se pueden aplicar en los títulos, resúmenes, palabras clave o metadatos y no necesariamente al texto completo del *paper*. Estas diferentes opciones estarán disponibles también en los distintos motores de búsqueda y serán utilizadas para refinar la obtención de información.

A priori, el método SMS no establece cuántas cadenas de búsqueda se deben generar. Para este trabajo, se ha considerado una única cadena de búsqueda con la siguiente estructura:

`<usability> AND <experimental study> AND software`

Como se observa, esta cadena está compuesta por tres elementos, concretamente, utilizando la palabra específica, es decir por tres *maxterms*. El último de estos términos suma consiste en una cadena simple de una sola palabra, por lo tanto, ni es necesario encerrarlo entre comillas ni necesita mucho más detalle de explicación.

Los otros dos términos tendrán la siguiente estructura interna:

- `<usability> = (usability OR accessibility OR "user experience" OR "user centered" OR "user centred")`
- `<experimental study> = ("empirical study" OR "empirical research" OR "empirical investigation" OR "empirical evaluation" OR "experimental study" OR "experimental research" OR "experimental investigation" OR "experimental evaluation")`

Llegado a este punto, nos centramos en seleccionar los motores de búsqueda en los que se procesará la cadena formulada anteriormente. Existen diversos motores de búsqueda de información científica. Por su relevancia y prestigio se han seleccionado los siguientes cinco: *IEEE Xplore*, *ACM Digital Library*, *Springer Link*, *ScienceDirect* y *Scopus*.

Destacar finalmente, que debido a las características particulares de cada uno de los motores de búsqueda, la cadena aplicada puede sufrir algún tipo de modificación para escoger únicamente los trabajos deseados. Por ejemplo, para *Scopus* habrá que solicitar que no se incluyan bases de datos de patentes pues se hará la búsqueda en cada una de ellas. De forma similar, en *ScienceDirect* se buscará sobre el grupo Ciencias Físicas e Ingeniería.

En la Tabla 3 se muestran los estudios encontrados en cada uno de los buscadores mencionados.

| Buscador | Encontrados |
|----------------------------|---------------|
| <i>IEEE Xplore</i> | 1.101 |
| <i>ACM Digital Library</i> | 4.298 |
| <i>Springer Link</i> | 263 |
| <i>ScienceDirect</i> | 6.009 |
| <i>Scopus</i> | 352 |
| TOTAL | 13.147 |

Tabla 3: Estudios encontrados en los cinco motores de búsqueda

2.4. Criterios de Inclusión-Exclusión

En este punto, ya se dispone de una selección de trabajos que se deben analizar. Como paso previo al análisis exhaustivo, se definen criterios de inclusión-exclusión que se

utilizarán para realizar distintos filtrados hasta obtener los estudios primarios con los que se contará para la finalización del método SMS.

Los criterios de inclusión definidos son los siguientes:

- A nivel de documento:
 - El documento está totalmente escrito en inglés.
 - El documento fue publicado en una revista, en un libro o capítulo de un libro, en un reporte técnico, en workshops o en una conferencia.
- En el ámbito del título:
 - El título se refiere a un estudio empírico, incluyendo explícitamente las siguientes palabras: “empirical”, “experimental”, “empirically” o “experimentally”.
 - El título se refiere a usabilidad.
- A nivel de contenidos:
 - El tema del estudio pertenece al área de usabilidad y se presenta al menos un estudio cuantitativo en el artículo.
 - Se hacen referencias a hipótesis.
 - Se utiliza estadística inferencial.
 - Se analiza el diseño del estudio.
 - Se analizan los resultados obtenidos.

Los criterios de exclusión determinados son los siguientes:

- En el ámbito del título:
 - El título hace referencia a una prueba de usabilidad o estudio de caso.
 - El título hace referencia a un meta-análisis.
 - El título hace referencia a una fase específica del estudio.
 - El título hace referencia a un concepto muy técnico de la usabilidad. Por ejemplo artículos que se refieren a estudios de usabilidad para los mandos de la videoconsola Wii.

La Tabla 4 muestra la secuencia de filtrado de los trabajos hasta llegar a los estudios primarios. Estos últimos se utilizarán en la siguiente sección para responder a las preguntas de investigación planteadas en la sección 2.2. Cabe destacar que a partir de los estudios encontrados con la cadena de búsqueda propuesta, los estudios candidatos se han ido filtrando de acuerdo a los criterios de inclusión-exclusión. Este proceso se ha realizado hasta que uno de los criterios fallaba momento en el que era descartado el trabajo. Además, indicar que en la columna Candidatos aparecen recogidos estudios duplicados mientras que en los estudios primarios se conservan en el primer buscador donde éstos se encuentren.

| Buscador | Encontrados | Candidatos | Estudios Primarios |
|----------------------------|---------------|------------|--------------------|
| <i>IEEE Xplore</i> | 1.101 | 76 | 1 |
| <i>ACM Digital Library</i> | 4.298 | 121 | 2 |
| <i>Springer Link</i> | 263 | 21 | 2 |
| <i>ScienceDirect</i> | 6.009 | 46 | 1 |
| <i>Scopus</i> | 352 | 48 | 1 |
| TOTAL | 13.147 | 312 | 7 |

Tabla 4: Detalle del filtrado realizado desde los estudios encontrados hasta obtener los estudios primarios

2.5. Resultados

En este punto ya podemos conseguir el producto del SMS, es decir, ya estamos en condiciones de responder a las cuestiones de investigación propuestas en la sección 2.2.

[Q1] ¿Cuáles son los temas, relacionados con la usabilidad, objeto de investigación empírica?

Hoy en día, el éxito de un sistema software depende fuertemente de la capacidad de uso de su interfaz. Por tanto, la evaluación de las interfaces se ha convertido en un aspecto crucial de la IS, [Cassino et al., 2015]. Evaluar una interfaz de usuario es el proceso por el que se determina el valor o la calidad de la misma en relación con unos objetivos marcados. No sólo es medir dicha calidad, sino que requiere de cierto componente crítico ya que con la evaluación se pretende alimentar el proceso de mejora continua de la interfaz de cara a conseguir que ésta sea lo más “usable” posible: el objetivo básico de la evaluación es conseguir el grado máximo de usabilidad. Se pueden realizar distintos tipos de evaluaciones: evaluación automática, evaluación empírica, evaluación formal y evaluación informal.

A su vez, cuando la evaluación de la usabilidad se realiza de forma previa a la distribución comercial de la interfaz existen distintos tipos de métodos: revisiones de expertos, test de usabilidad, test de laboratorio, encuestas, entrevistas y discusiones con usuarios.

Otro de los temas relevantes de estudio en usabilidad se presenta en [Fang & Holsapple, 2007]. Aquí se desarrolla una amplia experiencia donde se hace una síntesis de las dimensiones de la usabilidad web, incorporando detalles sobre la importancia de la psicología en esta rama.

Como ya se ha mencionado en el capítulo introductorio, la usabilidad es un factor de calidad del software que tiene como objetivo proporcionar la respuesta a muchos problemas encontrados en la interacción entre las personas y la tecnología. Al incorporar la usabilidad en aplicaciones web, denominada usabilidad web, es necesario refinar las definiciones generales para capturar la especificidad de esta clase de aplicaciones. En la web, las principales tareas que se realizan son: búsqueda directa de información deseada y servicios, descubrimiento de información a través de la navegación, comprensión de la información presentada. Por tanto, parafraseando la definición de ISO de la

usabilidad, la usabilidad web puede ser considerada como la capacidad de las aplicaciones web para apoyar este tipo de tareas con eficacia, eficiencia y satisfacción.

A su vez, y como se muestra en [Fang & Holsapple, 2007], la psicología tiene una gran importancia en esta rama de la IS. En [Horbæk & Frøkjær, 2004] se proponen dos tipos de técnicas de evaluación de la usabilidad basadas en la psicología. En este artículo se ratifica como los métodos de inspección clásicos, así como las heurísticas, no tienen explícitamente en cuenta los pensamientos de los sujetos en cada momento. Pero, sin embargo, al utilizar las dos técnicas que proponen (técnica cognitiva y metáforas del pensamiento del ser humano, MOT) consiguen una mejora en la inspección de la usabilidad obteniendo una que al utilizar MOT los participantes en el experimento son capaces de identificar muchos más problemas de usabilidad aunque es una técnica más difícil de comprender que la técnica cognitiva.

Un tercer tema de interés es la evaluación de la usabilidad desde aspectos sintácticos abordado en [Castro et al., 2011]. Aquí, como en las publicaciones anteriores, también se presenta una gran información acerca de la importancia de la usabilidad, pero este artículo no se centra en contenidos semánticos, tal y como se hace en [Fang & Holsapple, 2007] sino que se aborda la situación basándose en aspectos sintácticos, más relacionados con la forma de un sitio web concretando que una web debe ser capaz de guiar al usuario en los pasos que debe dar en la aplicación sin que éste pierda nunca la perspectiva de la tarea.

Por último, en [Bolchini, 2009] se presenta una perspectiva innovadora de la usabilidad: usabilidad como herramienta para mejorar los valores corporativos percibidos por los consumidores, que no debe confundirse con la percepción de la calidad de la marca que es, únicamente, un subconjunto del primer grupo.

[Q2] ¿Qué tipos de estudios cuantitativos se realizan en usabilidad?

En la actualidad existen diversos estudios relacionados con la usabilidad: estudios en los que se mide el impacto de la usabilidad a través de herramientas automáticas, estudios empíricos sobre usabilidad y cuasi-experimentos.

En [Cassino et al., 2015] se propone el uso de herramientas automáticas para la medición del impacto de la usabilidad. La utilización de este tipo de herramientas pueden reducir en gran medida los costes de las actividades tradicionales realizadas durante la evaluación por expertos o pruebas con usuarios para estimar la probabilidad de éxito de una aplicación. Sin embargo, es importante recalcar que estos métodos deben ser empíricamente válidos para demostrar su eficacia con respecto a los atributos que se supone que deben evaluar. En esta publicación se analizan cuatro sitios web y una aplicación interactiva independiente tanto de forma automática como con métodos manuales. El análisis de los resultados muestran que la metodología propuesta es un complemento útil a las técnicas estándar de evaluación de usabilidad.

El segundo tipo de estudios realizados en usabilidad son los estudios empíricos. Los estudios empíricos se basan en el método empírico que es un modelo de investigación científica que se sostiene sobre la experimentación y la lógica empírica. En [Piccioni et al., 2013] se propone el uso de estudios empíricos sobre usabilidad para crear bibliotecas API. En este trabajo se combinan entrevistas con preguntas basadas en el marco de dimensiones cognitivas con observaciones sistemáticas del comportamiento del programador mientras tiene que resolver tareas de programación descritas en fichas. El estudio proporcionó evidencias que demuestran que uno de los problemas más significativos es encontrar buenos nombres para las bibliotecas. Los 25 programadores participantes también demuestran dificultades para descubrir las relaciones entre los tipos de API. Se confirma la crucial importancia de la precisión de la documentación de la interfaz para conseguir mejora en la usabilidad.

En [Andreasen et al., 2007] se realiza un cuasi experimento en el que se evalúa la usabilidad de forma remota de tres formas distintas. Se determina que dichas técnicas funcionan mejor cuando se detectan problemas de usabilidad. Existen dos limitaciones importantes en este tipo de experimentos:

- (i) Son más vulnerables a los sesgos de selección, es decir, que el grupo de tratamiento puede diferir del grupo de control en características que están correlacionadas con los resultados estudiados, distorsionando los resultados del impacto.
- (ii) Dependen mucho de los métodos estadísticos multivariantes y son, por lo tanto, sensible al uso de modelos estadísticos apropiados y al tratamiento correcto de los problemas de estimación estadística.

Es necesario destacar en este punto, que en la bibliografía estudiada no se han encontrado estudios en los que se realicen experimentos verdaderos para evaluar si la presencia o ausencia de ciertos mecanismos de usabilidad afectan en la satisfacción o eficiencia de la aplicación web.

CAPÍTULO 3. EXPERIMENTO VERDADERO

3.1. Introducción

En este capítulo se presenta el experimento realizado en Galicia sobre la web *QuickStore* en noviembre de 2015. El objetivo del experimento ha sido estudiar el impacto de los mecanismos de usabilidad Preferencias (PRE), Abortar Operación (ABO) y Retroalimentación de Progreso (PRO), en los atributos de usabilidad eficiencia y satisfacción. Para ello, se ha realizado un estudio empírico con el fin de determinar si los usuarios de la herramienta web *QuickStore*, ver sección 3.7 y Anexo C, aprecian diferencias cuando disponen de alguno, o varios, de los mecanismos de usabilidad que están siendo evaluados.

Este capítulo se ha distribuido de la siguiente forma: en la sección 3.2 se detalla el diseño experimental definido. En la sección 3.3 se presentan los objetivos de investigación. A continuación, en la sección 3.4 se muestran las hipótesis de investigación. En la sección 3.5 se indican las variables respuesta que serán tenidas en cuenta y en la sección 3.6 se detalla el perfil de los sujetos que han intervenido en el experimento. En la sección 3.7 se detalla la instrumentación, es decir, aquí se describen las tres tareas que se deben realizar sobre la aplicación web *QuickStore* en la que se ha realizado el experimento. En la sección 3.8 se indica el método de recolección de datos. En la sección 3.9 se describen las amenazas a la validez interna y externa del experimento. La sección 3.10 introduce el análisis de datos. Por último, en las secciones 3.11 y 3.12 se presentan los resultados estadísticos descriptivos y el test ANOVA, respectivamente.

3.2. Diseño Experimental

Como en casi todos los procesos de investigación experimental, el diseño del experimento sufre modificaciones desde la etapa inicial hasta la puesta en marcha del experimento final. En esta sección, se detalla el diseño experimental utilizado en el experimento verdadero final de esta investigación. En el Anexo A se muestra la evolución de los diseños durante todo el proyecto de investigación, incluyendo el diseño utilizado durante el experimento piloto.

El diseño del experimento verdadero es un diseño *between subjects*. En un diseño *between subjects*, cada sujeto experimental está asignado únicamente a una combinación de los niveles de los factores. En otras palabras, cada sujeto experimental proporciona un único valor para cada variable respuesta que será utilizado en el análisis estadístico. En el caso particular de nuestro estudio cada sujeto ha participado en varios experimentos (uno por cada mecanismo de usabilidad considerado), contrabalanceando el orden de ejecución de las tareas (una por mecanismo) para anular posibles efectos de presentación de las tareas. Este diseño está caracterizado por la matriz de grupos presentada en la Tabla 5 (los ceros indican que el mecanismo de usabilidad no está

presente mientras que, por el contrario, los unos indican que el mecanismo sí está presente) y por la matriz con los órdenes de exposición de los mecanismos de usabilidad, que representan las secuencias posibles de realización de las tres tareas asociadas a cada uno de los mecanismos, descritos en la Tabla 6. El diseño se completa con la matriz de asignación de agrupaciones que se describe en la Tabla 7.

| | ABO | PRO | PRE |
|---|-----|-----|-----|
| A | 0 | 0 | 1 |
| B | 1 | 0 | 0 |
| C | 0 | 1 | 0 |
| D | 1 | 1 | 1 |

Tabla 5: Matriz de grupos

| | Primero | Segundo | Tercero |
|----|---------|---------|---------|
| O1 | ABO | PRO | PRE |
| O2 | ABO | PRE | PRO |
| O3 | PRO | PRE | ABO |
| O4 | PRO | ABO | PRE |
| O5 | PRE | ABO | PRO |
| O6 | PRE | PRO | ABO |

Tabla 6: Matriz con el orden de exposición de cada mecanismo de usabilidad

| | | | Primero | | | Segundo | | | Tercero | | |
|---|--------|----|---------|-----|-----|---------|-----|-----|---------|-----|-----|
| | | | ABO | PRO | PRE | ABO | PRO | PRE | ABO | PRO | PRE |
| A | AGR 1 | O1 | 0 | | | | 0 | | | | 1 |
| | AGR 2 | O2 | 0 | | | | | 1 | | 0 | |
| | AGR 3 | O3 | | 0 | | | | 1 | 0 | | |
| | AGR 4 | O4 | | 0 | | 0 | | | | | 1 |
| | AGR 5 | O5 | | | 1 | 0 | | | | 0 | |
| | AGR 6 | O6 | | | 1 | | 0 | | 0 | | |
| B | AGR 7 | O1 | 1 | | | | 0 | | | | 0 |
| | AGR 8 | O2 | 1 | | | | | 0 | | 0 | |
| | AGR 9 | O3 | | 0 | | | | 0 | 1 | | |
| | AGR 10 | O4 | | 0 | | 1 | | | | | 0 |
| | AGR 11 | O5 | | | 0 | 1 | | | | 0 | |
| | AGR 12 | O6 | | | 0 | | 0 | | 1 | | |
| C | AGR 13 | O1 | 0 | | | | 1 | | | | 0 |
| | AGR 14 | O2 | 0 | | | | | 0 | | 1 | |
| | AGR 15 | O3 | | 1 | | | | 0 | 0 | | |
| | AGR 16 | O4 | | 1 | | 0 | | | | | 0 |
| | AGR 17 | O5 | | | 0 | 0 | | | | 1 | |
| | AGR 18 | O6 | | | 0 | | 1 | | 0 | | |
| D | AGR 19 | O1 | 1 | | | | 1 | | | | 1 |
| | AGR 20 | O2 | 1 | | | | | 1 | | 1 | |
| | AGR 21 | O3 | | 1 | | | | 1 | 1 | | |
| | AGR 22 | O4 | | 1 | | 1 | | | 1 | | 1 |
| | AGR 23 | O5 | | | 1 | 1 | | | | 1 | |
| | AGR 24 | O6 | | | 1 | | 1 | | 1 | | |

Tabla 7: Matriz de asignación de agrupaciones con órdenes de exposición

3.3. Objetivos de Investigación

El propósito de este experimento, tal y como se ha introducido en la sección 1.1, es resolver la siguiente pregunta de investigación:

*¿La presencia de los mecanismos de usabilidad
impacta en la usabilidad de la aplicación?*

Más concretamente, se pretenden responder a las siguientes tres preguntas de investigación:

- I. ¿La presencia del mecanismo de usabilidad Abortar Operación impacta en la usabilidad de la aplicación?
- II. ¿La presencia del mecanismo de usabilidad Retroalimentación de Progreso impacta en la usabilidad de la aplicación?
- III. ¿La presencia del mecanismo de usabilidad Preferencias impacta en la usabilidad de la aplicación?

En el caso de este experimento verdadero, la herramienta utilizada será la aplicación *QuickStore*, introducida en el Anexo C, incorporando las modificaciones pertinentes fruto de las conclusiones del experimento piloto detallado en el Anexo B.

3.4. Hipótesis de Investigación

De acuerdo con los objetivos de investigación formulados y considerando que la variable respuesta usabilidad se descompone en las variables eficiencia y satisfacción, las hipótesis de investigación que se establecen son las mostradas en la siguiente Tabla 8.

| Variable R | MU | Hipótesis | Descripción |
|------------------|--|----------------------|--|
| 1. Eficiencia | 1. Abortar Operación (ABO) | H.1.1.0 (nula) | No existe diferencia significativa en la eficiencia del usuario al incorporar el MU-ABO o al no incorporarlo. |
| | | H.1.1.1 (alterna) | Existe diferencia significativa en la eficiencia del usuario al incorporar el MU-ABO o al no incorporarlo. |
| | 2. Retroalimentación de Progreso (PRO) | H.1.2.0 (nula) | No existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRO o al no incorporarlo. |
| | | H.1.2.1 (alterna) | Existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRO o al no incorporarlo. |
| | 3. Preferencias (PRE) | H.1.3.0 (nula) | No existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRE o al no incorporarlo. |
| | | H.1.3.1 (alterna) | Existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRE o al no incorporarlo. |

| Variable R | MU | Hipótesis | Descripción |
|--------------------|---|----------------------|--|
| 2. Satisfacción | 1. Abortar Operación (ABO) | H.2.1.0 (nula) | No existe diferencia significativa en la satisfacción del usuario al incorporar el MU-ABO o al no incorporarlo. |
| | | H.2.1.1 (alterna) | Existe diferencia significativa en la satisfacción del usuario al incorporar el MU-ABO o al no incorporarlo. |
| | 2. Retroalimentación de Progreso (PRO) | H.2.2.0 (nula) | No existe diferencia significativa en la satisfacción del usuario al incorporar el MU-PRO o al no incorporarlo. |
| | | H.2.2.1 (alterna) | Existe diferencia significativa en la satisfacción del usuario al incorporar el MU-PRO o al no incorporarlo. |
| | 3. Preferencias (PRE) | H.2.3.0 (nula) | No existe diferencia significativa en la satisfacción del usuario al incorporar el MU-PRE o al no incorporarlo. |
| | | H.2.3.1 (alterna) | Existe diferencia significativa en la satisfacción del usuario al incorporar el MU-PRE o al no incorporarlo. |

Tabla 8: Hipótesis de investigación

3.5. Variables Respuesta

En este experimento se utilizarán dos variables respuesta clásicas en los experimentos de usabilidad: la eficiencia y la satisfacción.

En el caso de la *eficiencia* se medirán dos parámetros: el tiempo medido en segundos que un sujeto necesita para completar la tarea, es decir la eficiencia como medida de la rapidez [Tullis & Stetson, 2004] y el número de clicks que un sujeto necesita para completar la tarea, o dicho de otra forma, el nivel de interacción con el sistema para cada uno de los usuarios [Frøkjær et al., 2000].

La *satisfacción* se medirá con el valor medio de las respuestas a las preguntas de los cuestionarios post-tareas diseñados. Las preguntas tienen valores ordinales de la escala de Likert (uno = totalmente en desacuerdo al cinco = totalmente de acuerdo) [Sauro & Kindlund, 2005].

3.6. Sujetos

Los participantes, o sujetos experimentales, están representados por estudiantes que han participado de forma voluntaria y con consentimiento de los profesores del Instituto de Rodeira, Galicia, involucrados en el proyecto y que tendrán que realizar las distintas tareas propuestas durante el experimento.

Con el fin de evitar el efecto de aprendizaje, no se ha elaborado ningún tipo de sesión previa de información o de práctica con los propios participantes. Para la realización del experimento, simplemente, se ha proporcionado una información básica para que los participantes tengan clara cuál es la estructura del experimento y sean capaces de

realizar las tareas del mismo, así como contestar coherentemente las preguntas de los cuestionarios.

Por último, todos los participantes serán asignados al azar a una de las veinticuatro posibles agrupaciones del diseño, ver sección 3.2. La agrupación será la que determine qué mecanismos de usabilidad estarán disponibles en el experimento y cuáles no.

A partir del análisis del cuestionario de familiaridad, que todos los participantes han cubierto como primer paso del experimento, destacan las siguientes características: la muestra cuenta con 100 sujetos de los cuales, 43 son hombres y 57 mujeres. En cuanto a la edad, 85 de ellos están en la franja 16-18 años, 10 pertenecen al intervalo 18-30 y, únicamente, 5 de ellos son mayores de 30 años.

Por las edades de los sujetos claramente la ocupación mayoritaria, 90, es estudiante. Concretamente, son alumnos de 1º y 2º curso de Bachillerato. Nueve de ellos ya han iniciado su vida laboral, aunque son también estudiantes pues están cursando el Ciclo Superior de Electrónica. También ha participado un profesor en el experimento.

De forma completamente mayoritaria, los sujetos hacen uso de Internet en sus casas. Algunos de ellos lo utilizan también en el trabajo o en otros ámbitos. El uso principal es entretenimiento y educación. En este mismo cuestionario se pregunta por las costumbres de los sujetos referidas a las compras en Internet. Estas tendencias se ven reflejadas en la Tabla 9.

| <i>¿Has comprado en Internet personalmente?</i> | | | | |
|---|----------------|---------------|-------------------|--------------|
| Nunca 37 | Rara Vez 17 | A Veces 32 | Casi Siempre 9 | Siempre 5 |

Tabla 9: Información referente a compras

3.7. Instrumentación

Para realizar el experimento, como ya se ha comentado, se dispone de la web *QuickStore*, <http://webadm.senado.gov.py/tesisweb/>, que sigue el diseño detallado en la sección 3.2. En esta sección se comentarán cada una de las tareas que deben hacer los sujetos de experimentación para completar su participación. Puede consultarse el Anexo C para comprobar la apariencia de la web *QuickStore* y para poder ver paso a paso cómo resolver las cuatro tareas presentadas a continuación.

1. *Tarea Abortar Operación* (Tarea ABO): el usuario simulará una operación de cancelar durante el procesamiento de su carrito de compra. Al ingresar al sistema, el carrito del usuario contendrá varios artículos. La tarea consiste en que el sujeto ingrese a su carrito de compra y modifique los datos del carro (por ejemplo, incrementar la cantidad de cualquiera de los artículos, completar un código promocional, etc.) y, seguidamente, cancele la operación. Si el mecanismo de usabilidad está presente, el usuario dispondrá de un botón *Cancelar*, el cual, únicamente, solicitará confirmación sobre los cambios pendientes. Por el contrario, si el mecanismo de usabilidad no está activo la

única forma de realizar la cancelación es la manual, es decir, debe deshacer los cambios que realizó desde el inicio de la tarea.

2. *Tarea Retroalimentación de Progreso* (Tarea PRO): en esta segunda tarea, el usuario simulará una operación de búsqueda de un artículo concreto y deberá agregarlo al carrito de compra una vez lo localice. El sujeto inicia la tarea desde la pantalla principal de la herramienta *QuickStore*, en la que puede realizar una búsqueda con los criterios que estime, por ejemplo, el nombre de la novela que se propone. Si la búsqueda es exitosa, únicamente debe pulsar en el botón *Añadir al Carrito*.
Si el mecanismo de usabilidad está activo, mientras se realiza la búsqueda se mostrará una barra de progreso informando al usuario de que la acción se está ejecutando con un mensaje, al final, indicando la cantidad de artículos encontrados.
Si el mecanismo de usabilidad no está presente, se ejecutará igualmente la búsqueda, pero no se informará al usuario de que la acción se está realizando.
3. *Tarea Preferencias* (Tarea PRE): esta tarea está dividida en dos partes. En primer lugar se realizará la tarea preferencias básica y, seguidamente, la tarea preferencias ficticia. Se detallan ambas a continuación.
 - a. *Tarea Preferencias Normal o Básica* (Tarea PRE-B): si el mecanismo de usabilidad está activo, el usuario podrá personalizar ciertos aspectos de la tienda a su gusto. Inicialmente, el aspecto de la interfaz no será agradable, pues el tamaño de la letra será pequeño y la tipografía será poco legible. Por el contrario, si el mecanismo no está presente el usuario no podrá modificar la apariencia de la interfaz desde la propia aplicación.
 - b. *Tarea Preferencias Ficticia* (Tarea PRE-F): en este caso el usuario deberá encontrar cierta información, proporcionada en la aplicación, relacionada con los plazos de devolución de los productos comprados. Si el usuario ha realizado la modificación de la interfaz del sistema podrá ver fácilmente dónde está el enlace para encontrar la información que se propone. Sin embargo, si el usuario no ha sido capaz, o no ha querido, modificar la apariencia de la aplicación, tendrá más difícil encontrar la información requerida. Para esta tarea únicamente se considera la eficiencia y la satisfacción del usuario.

El conjunto de estas cuatro tareas conforman el escenario de ejecución del experimento. La selección de las tareas para cada mecanismo de usabilidad propone que el usuario realice tareas como cliente de la tienda, es decir, interactúa con la interfaz final del sistema haciendo tareas no administrativas.

Además, a cada tarea (ABO, PRE, PRO) se le asocia un cuestionario post-tarea donde se recolectará información sobre una de las variables respuesta que se investigan: la satisfacción de los usuarios durante la realización de la tarea. Para medir la eficiencia se registrará el tiempo y el número de clicks que realizan los sujetos durante el experimento.

Por último, destacar que los sujetos deben cumplimentar el cuestionario de familiaridad de forma previa a la realización de las tareas propuestas. Este cuestionario se muestra en el Anexo F.

3.8. Procedimiento de Recolección de Datos

El experimento se ha realizado en dos sesiones planificadas el 6/11/2015 y el 9/11/2015 durante las cuales se han recolectado los siguientes datos:

1. Datos referidos al perfil de cada uno de los sujetos.
2. Tiempo utilizado para resolver cada una de las tareas propuestas.
3. Número de clicks utilizados para resolver cada una de las tareas propuestas.
4. Valores indicados por los usuarios en los cuestionarios propuestos después de cada una de las tareas.

3.8.1. Métricas

A continuación, se muestran, en la Tabla 10, los datos de métricas -tiempo y número de clicks- recolectados. Resaltar que se obtiene una tabla por cada uno de los mecanismos de usabilidad y ésta tiene siempre la misma estructura.

| MU | Info Tarea | Sujeto | Agrupación | IP | Fecha Hora | Tiempo | Clicks |
|----|------------|--------|------------|----|------------|--------|--------|
| | | | | | | | |

Tabla 10: Columnas con los datos de las métricas observadas

- *MU*: mecanismo de usabilidad.
- *Info Tarea*: se utiliza en el mecanismo Preferencias para discernir entre la tarea normal y la tarea ficticia.
- *Sujeto*: identificador que introduce el participante en el cuestionario de familiaridad al inicio del experimento.
- *Agrupación*: cada una de las 24 agrupaciones disponibles.
- *IP*: IP utilizada durante el experimento.
- *Fecha Hora*: fecha y hora de realización del experimento.
- *Tiempo*: tiempo utilizado en la realización de la tarea del mecanismo de usabilidad del que se está recolectando la información.
- *Clicks*: número de clicks utilizados en la realización de la tarea del mecanismo de usabilidad del que se está recolectando la información.

En la Tabla 11, y de forma coherentemente, con el número de sujetos participantes se pueden observar los números de filas de las métricas.

| Tabla de Métrica | Tipo de Tarea | Número de Filas |
|---|----------------|-----------------|
| Tabla métrica Abortar Operación | Tarea Normal | 100 |
| Tabla métrica Preferencias | Tarea Normal | 100 |
| | Tarea Ficticia | 100 |
| Tabla métrica Retroalimentación de Progreso | Tarea Normal | 100 |

Tabla 11: Número de filas en las tablas de métricas

3.8.2. Cuestionarios

A continuación, se muestran los datos de los cuestionarios recolectados -cuestionario de familiaridad y cuestionarios asociados a cada una de las tareas-. En este caso se obtienen cinco tablas de datos diferentes: una para almacenar la información de los cuestionarios de familiaridad y tres para los datos asociados a cada una de las tareas, una por cada mecanismo de usabilidad. Sin embargo, solamente hay dos estructuras distintas para estas tablas: la referida al cuestionario de familiaridad y la tabla para el resto de cuestionarios (Tabla 12).

Cuestionario de familiaridad:

En este cuestionario se pregunta la edad de los sujetos, su ocupación y el sexo. También se preguntan los idiomas conocidos por el usuario y si es informático o no. Por último, se hacen preguntas acerca de cómo utilizan los sujetos Internet. Concretamente las cuestiones son:

- ¿Dónde usas Internet?
- ¿Para qué usas Internet?
- ¿Qué aplicaciones usas?
- ¿Has comprado personalmente en Internet?
- ¿Has comprado con intermediarios?
- ¿Cuáles son tus tiendas preferidas?
- ¿Qué problemas has percibido en las compras?
- ¿Consideras atractivas las compras por Internet?

Cuestionarios propios de las tareas:

| MU | Id Pregunta | Pregunta | Sujeto | Agrupación | Fecha Hora | IP | Valor | Comentario |
|----|-------------|----------|--------|------------|------------|----|-------|------------|
| | | | | | | | | |

Tabla 12: Columnas de las tablas con datos de los cuestionarios

- *MU*: mecanismo de usabilidad.
 - ✓ *Abortar Operación*
 - ✓ *Preferencias*
 - ✓ *Retroalimentación de Progreso*
- *Id Pregunta*: identificador de la pregunta.
- *Pregunta*: texto de la pregunta.
- *Sujeto*: identificador que introduce el participante en el cuestionario de familiaridad al inicio del experimento.
- *Agrupación*: cada una de las 24 agrupaciones disponibles.
- *Fecha Hora*: fecha y hora de realización del experimento.
- *IP*: IP utilizada durante el experimento.
- *Valor*: puntuación obtenida entre las siguientes:
 - ✓ 1 – Totalmente en desacuerdo.
 - ✓ 2 – En desacuerdo.
 - ✓ 3 – Neutral.
 - ✓ 4 – De acuerdo.
 - ✓ 5 – Totalmente de acuerdo.
- *Comentario*: comentario escrito por el sujeto para matizar la información proporcionada.

En este caso, cada una de las tablas contiene la siguiente cantidad de información:

| Tabla de Cuestionarios | Número de Filas |
|--|-----------------|
| Tabla Cuestionario Familiaridad | 100 |
| Tabla Cuestionario Abortar Operación | 300 |
| Tabla Cuestionario Preferencias | 400 |
| Tabla Cuestionario Retroalimentación de Progreso | 300 |

Tabla 13: Número de filas de las tablas con información de los cuestionarios

3.9. Validez Interna y Externa

En esta sección se comentan todos los aspectos relacionados con la validez del experimento que se presenta.

La validez interna se articula en torno a la supervisión del proceso en aras de establecer las relaciones de asociación entre las variables independientes y la dependiente. La validez externa está relacionada con el establecimiento de las condiciones que permiten la generalización de los resultados al ámbito natural en el que aparecen los procesos investigados [Acuña et al., 2008].

3.9.1. Validez Interna

Tal y como se ha introducido, la validez interna se refiere a cuánta confianza tenemos en que los resultados del experimento sean posibles interpretarlos y éstos sean válidos. Esta validez se relaciona con la calidad del experimento y se logra cuando hay control, cuando los grupos difieren entre sí solamente en la exposición a la variable independiente, cuando las mediciones de la variable dependiente son fiables y válidas y cuando el análisis es el adecuado para el tipo de datos que se manejan [Campbell & Stanley, 1963].

En este experimento existen varias potenciales amenazas a la validez:

1. *Conocimiento de la tecnología:* aunque todos los participantes parten de un mismo nivel de experiencia hacia este tipo de experimentos (nivel novato), no todos los sujetos tienen el mismo nivel de conocimiento de la actividad que se va a desarrollar. Del cuestionario de familiaridad se desprende que hay diferencias en el uso de páginas o aplicaciones web para la realización de compras online.
2. *Orden de realización de las tareas:* podría producir un sesgo por efecto de aprendizaje.
3. *Baja experiencia de los usuarios:* como se ha comentado en la sección 3.6, todos los sujetos que han participado en este experimento eran primerizos y esto puede llevar a que no entiendan bien cómo realizarlo, no presten especial atención a las indicaciones previas, etc.

Las dos primeras amenazas se han mitigado asignando al azar a los sujetos a cada agrupación. [Cochran & Cox, 1980] mencionan que “la aleatorización es en cierta forma análoga a un seguro por el hecho de que es una precaución contra interferencias que pueden o no ocurrir, y ser o no importantes, si ocurren. Generalmente, es aconsejable tomarse el trabajo de aleatorizar, aún cuando no se espere que haya un sesgo importante al dejar de hacerlo”.

La tercera amenaza se suple introduciendo el orden como un factor dentro del diseño tal y como se establece en la sección 3.2.

3.9.2. Validez Externa

La validez externa consiste en el poder de generalización de los resultados obtenidos. En este experimento los resultados no son generalizables a toda la población de usuarios, pero los sujetos participantes en sí mismos generan bastante confianza porque no son informáticos y no están sesgados por la tecnología ya que son jóvenes familiarizados con la tecnología a nivel de usuario (todos los sujetos son usuarios asiduos de Internet y son capaces de enfrentarse a este tipo de herramientas interactivas con eficacia).

Por ello, estos sujetos son personas que conforman una parte importante de la población de usuarios que hace uso de aplicaciones web de tienda para la compra de productos y que permiten obtener evidencias empíricas sobre el impacto de los mecanismos de usabilidad analizados a nivel de usuarios no informáticos avanzados.

3.10. Análisis de Datos

El método estadístico utilizado para la realización del análisis de este experimento se puede resumir con el esquema de la Figura 3:

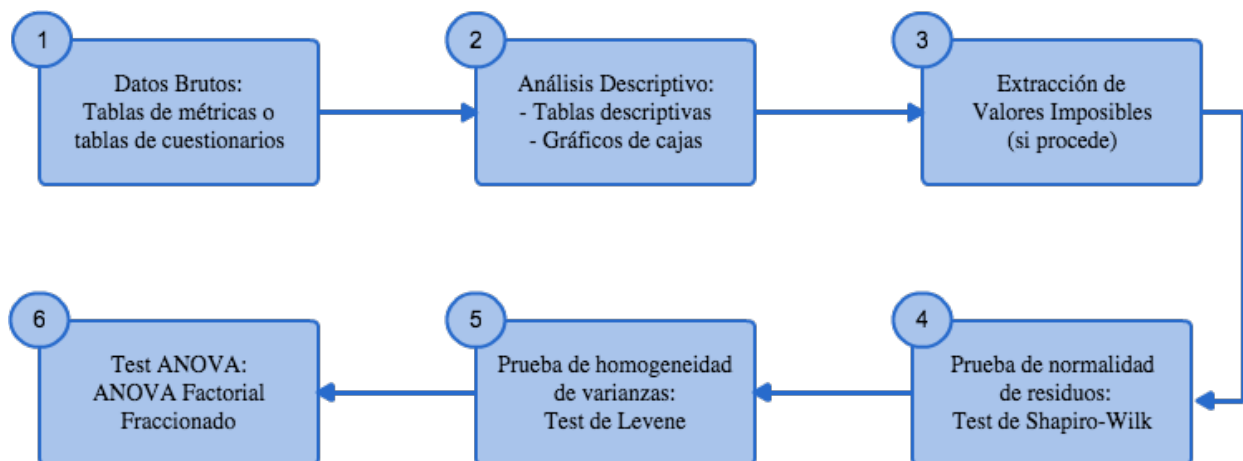


Figura 3: Método estadístico

1. *Datos Brutos*: datos contenidos en las tablas de métricas bajo las indicaciones detalladas en la sección 3.8 de este capítulo, en el caso de la eficiencia. Datos

contenidos en las tablas de los cuestionarios de las tareas en el caso de la satisfacción.

2. *Análisis descriptivo*: como primer paso para entender los datos, se efectuará un análisis básico descriptivo. En este caso, se han detectado datos irregulares, razón por la cual se han decidido eliminar, tal y como se ha comentado en el paso anterior. La importancia de este análisis radica en que informa de cómo es la muestra de la que se dispone. Entre otros, se calcularán los siguientes estadísticos descriptivos: mínimo, máximo, media, mediana, desviación típica, coeficiente de asimetría, rango intercuartílico, etc. Para completar este análisis descriptivo se realizarán gráficos de caja que ayuden a comprender la forma de los datos.
3. *Extracción de valores imposibles*: tras analizar los estadísticos descriptivos y para mantener la coherencia de los resultados se han decidido eliminar los valores imposibles. Se consideran imposibles aquellos valores que por las características del experimento determinan que la tarea o bien no se ha hecho o bien se ha hecho de forma incorrecta, por ejemplo, eliminar aquellos sujetos que han realizado una tarea con 0 clicks o en 0 segundos.

En cuanto a los datos de satisfacción, no tiene sentido eliminar valores imposibles, pues los valores siempre están controlados por la mínima y máxima puntuación posible en los cuestionarios.

4. *Prueba de normalidad*: para continuar el análisis se realiza la prueba de normalidad de residuos. Concretamente, se realiza una de las pruebas más utilizadas para este fin, la prueba de *Shapiro-Wilk*.
5. *Prueba de homogeneidad*: por último, previamente a la realización del test ANOVA es necesario asegurar la homogeneidad de las varianzas. Para ello, se utiliza el test de *Levene*, un test estadístico utilizado para evaluar la igualdad de las varianzas para una variable calculada para dos o más grupos.
6. *Test ANOVA*: por último lugar, se realizará un test ANOVA para comprobar si se puede aceptar o rechazar la hipótesis nula en cada caso. Concretamente, se realizará un test ANOVA factorial fraccionado. La Tabla 14 muestra los distintos elementos considerados en el test ANOVA, independientemente del mecanismo de usabilidad que se pretenda analizar. Se puede observar en la misma como existen tres factores: un factor principal que es cada mecanismo de usabilidad con dos niveles presencia y ausencia, un segundo factor correspondiente al grupo con cuatro niveles denominados A, B, C y D y un tercer factor el orden de realización de las tareas con seis niveles (Orden 1 - ... - Orden 6).

| Factor | Grado | Valores posibles |
|--------|-----------|-----------------------|
| PA | Principal | Presencia - Ausencia |
| Grupo | Segundo | A – B – C – D |
| Orden | Tercero | 1 – 2 – 3 – 4 – 5 – 6 |

Tabla 14: Elementos considerados en los test ANOVA

3.11. Análisis Descriptivo

Como se ha explicado en la sección 3.10 se realiza un análisis descriptivo para conocer la estructura de las muestras obtenidas, analizando los siguientes parámetros estadísticos:

Mínimo – Máximo – Media – Mediana – Rango Intercuartílico – Primer Cuartil – Segundo Cuartil – Tercer Cuartil – Cuarto Cuartil – Desviación Típica – Varianza – Coeficiente de Skewness – Coeficiente de Curtosis.

El análisis se realizará, para cada variable respuesta, comparando los resultados obtenidos por aquellos sujetos que tenían presente (P) el mecanismo de usabilidad contra los usuarios que tenían el mecanismo ausente (A). En el Anexo D se pueden encontrar los diagramas de caja que representan, además del factor Presencia/Ausencia, el factor Orden y el factor Grupo.

3.11.1. Eficiencia

En esta subsección se exponen tanto los resultados obtenidos en el número de clicks como el tiempo, en segundos, las dos medidas que, como hemos comentado, nos sirven para medir la eficiencia de los sujetos en cada una de las tareas.

3.11.1.1. Preferencias

Como se muestra en la Tabla 15 la media de clicks es inferior en el caso de la presencia del mecanismo Preferencias frente a la ausencia de este mecanismo. Además, la desviación típica es ligeramente menor en el primer caso que en el segundo. Sin embargo, destaca la diferencia entre los valores máximos: en el caso de presencia del mecanismo la persona que más clicks empleó superó en nueve a la persona que más clicks utilizó en el caso de la ausencia del mecanismo Preferencias. En cuanto a la asimetría, se observa que en ambos casos es positiva, aunque mayor en el caso de presencia, lo que supone que la mayoría de los datos superan el valor de la media. Además, observando el coeficiente de Curtosis se puede comprobar como, en ambos casos, la muestra es leptocúrtica aunque, nuevamente, es mayor esta tendencia en el caso de la presencia del mecanismo.

En el caso del tiempo, la tendencia es distinta. En media, las personas que disponían del mecanismo de usabilidad tardaron más (del orden de medio minuto) que aquellos usuarios que no disponían del mecanismo. La desviación típica es menor en el caso de la presencia del mecanismo frente a la ausencia. Es importante también que los coeficientes de asimetría y Curtosis tomen valores positivos, pero más moderados que

en el caso del estudio del número de clicks. Aquí, la asimetría es menor en el caso de la presencia del mecanismo y mayor (casi el doble) en el caso de la ausencia del mecanismo, pero como se decía, en ambos casos, es positiva. Existe un comportamiento similar a la hora de estudiar el coeficiente de Curtosis.

| PRE | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|----------|------|--------|--------|---------|--------|-------|--------|---------|--------|-------|----------|----------|----------|
| Clicks P | 0 | 54 | 7.92 | 4 | 9 | 2 | 4 | 11 | 54 | 9.83 | 96.81 | 2.48 | 8.10 |
| Clicks A | 0 | 45 | 9.5 | 6 | 8 | 2 | 6 | 10 | 45 | 11.80 | 139.18 | 1.89 | 2.70 |
| Tiempo P | 1.69 | 373.21 | 121.55 | 113.36 | 114.26 | 51.88 | 113.36 | 166.139 | 373.21 | 88.16 | 7771.33 | 0.96 | 0.55 |
| Tiempo A | 5.53 | 432.16 | 112.70 | 89,60 | 105.81 | 41.82 | 89.60 | 147.62 | 432.16 | 99.28 | 9856.28 | 1.80 | 3.28 |

Tabla 15: Estadísticos descriptivos para el mecanismo Preferencias en el estudio de Eficiencia

Se completa el análisis de los estadísticos descriptivos mostrando los diagramas de caja para cada una de las dos variables respuesta, número de clicks y tiempo, Figura 4 y Figura 5 respectivamente.

En el caso del número de clicks, se observa que el comportamiento de los sujetos con ausencia del mecanismo es bastante simétrico y tiene cinco valores atípicos. Sin embargo, en el caso de la presencia del mecanismo la muestra no es simétrica, la mediana es más baja que en el caso de la ausencia, pero el tercer cuartil es mayor que en el caso de la ausencia. Aquí existen cuatro valores atípicos.

En el caso del tiempo, se observa un comportamiento bastante similar entre ambas muestras aunque tomando valores un poco superiores de Q1, mediana y Q3 en el caso de la presencia del mecanismo. Además, existen cinco valores atípicos, tres de ellos en la muestra Ausencia y los dos restantes en la Presencia.

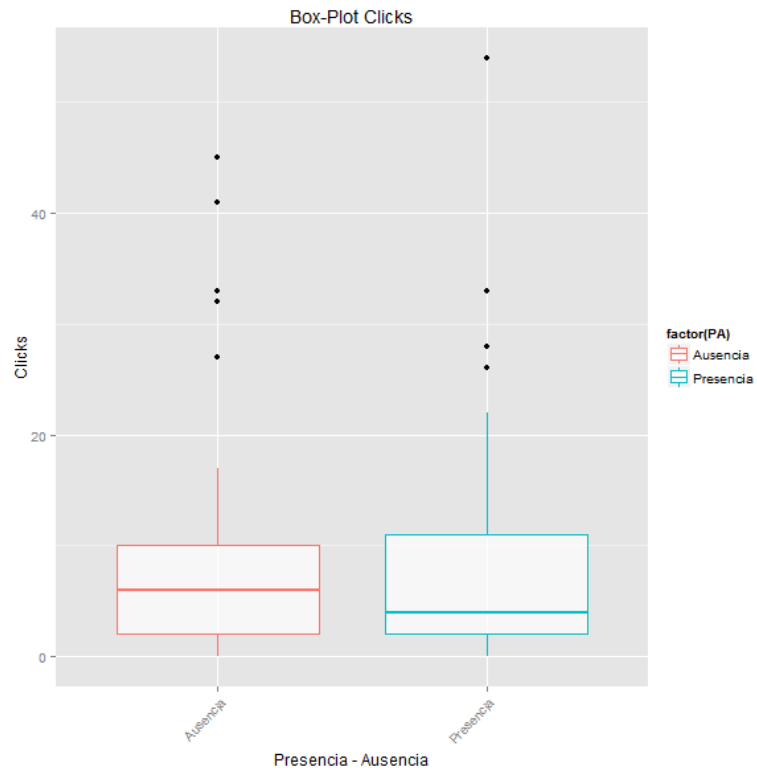


Figura 4: Diagrama de cajas para el Número de Clicks en el mecanismo Preferencias

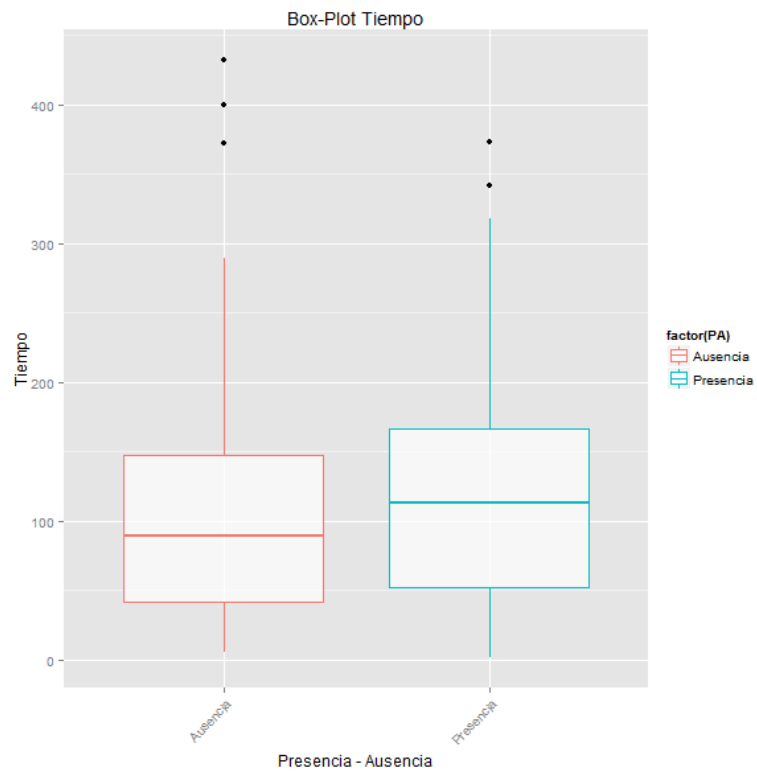


Figura 5: Diagrama de cajas para el Tiempo, en segundos, en el mecanismo Preferencias

3.11.1.2. Retroalimentación de Progreso

La Tabla 16 muestra los estadísticos descriptivos calculados para el mecanismo Retroalimentación de Progreso en el caso de la eficiencia.

En el caso del número de clicks la media es inferior en el caso de la presencia del mecanismo, con una variación típica menor aunque con un coeficiente de asimetría y Curtosis positivo y mayor. Destaca que en el caso de la presencia del mecanismo coinciden los dos primeros cuartiles. En el caso de la ausencia del mecanismo, la desviación típica toma un valor de 4.40 y el coeficiente Curtosis, aunque positivo, está próximo a cero, esto supone que la muestra es bastante similar a una muestra normal.

En el caso de la variable tiempo se observa que los valores son superiores en el caso de la ausencia del mecanismo para las variables media, mediana, cuartiles, varianza, coeficiente de asimetría y coeficiente de Curtosis (en ambos casos positivos). Sin embargo, la desviación típica es significativamente inferior en el caso de la ausencia del mecanismo frente a la presencia.

| PRO | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|----------|-------|-------|-------|---------|-------|-------|-------|-------|--------|-------|----------|----------|----------|
| Clicks P | 0 | 15 | 4.62 | 3 | 3 | 3 | 3 | 6 | 15 | 2.80 | 7.85 | 1.83 | 4.35 |
| Clicks A | 0 | 19 | 7.39 | 6 | 4 | 5 | 6 | 9 | 19 | 4.40 | 19.42 | 1.16 | 0.55 |
| Tiempo P | 23.20 | 194 | 73.36 | 67.78 | 37.05 | 46.83 | 67.78 | 83.87 | 193.96 | 35.75 | 1278.26 | 1.11 | 1.27 |
| Tiempo A | 9.39 | 227.3 | 77.36 | 68.012 | 36.57 | 50.22 | 68.01 | 86.80 | 227.26 | 10.19 | 1615.88 | 1.41 | 3.021 |

Tabla 16: Estadísticos descriptivos para el mecanismo Retroalimentación de Progreso en el estudio de Eficiencia

Los diagramas de caja completan, de forma gráfica, la información sobre los estadísticos descriptivos. Permiten ver claramente como en el caso de la variable número de clicks, Figura 6, no tiene una muestra simétrica ni en el caso de la presencia del mecanismo ni en el caso de la ausencia del mecanismo. Se observa también como los valores de los cuartiles son superiores en el caso de la ausencia del mecanismo respecto a la presencia del mismo. En cuanto al número de *outliers* es parecido en ambos casos: en presencia hay dos valores atípicos, frente a los tres que hay en el caso de la ausencia.

Estudiando ahora la variable tiempo, Figura 7, se observa una tendencia muy parecida en los dos grupos diferenciados. Tanto los cuartiles primero y tercero, como la mediana (segundo cuartil), toman valores muy similares en las dos situaciones. Coincide también el número de *outliers* con tres valores atípicos en cada caso. La diferencia más significativa se encuentra en el bigote inferior de las cajas: en el caso de la presencia del mecanismo, este bigote es más corto que en el caso de la ausencia del mecanismo Retroalimentación de Progreso.

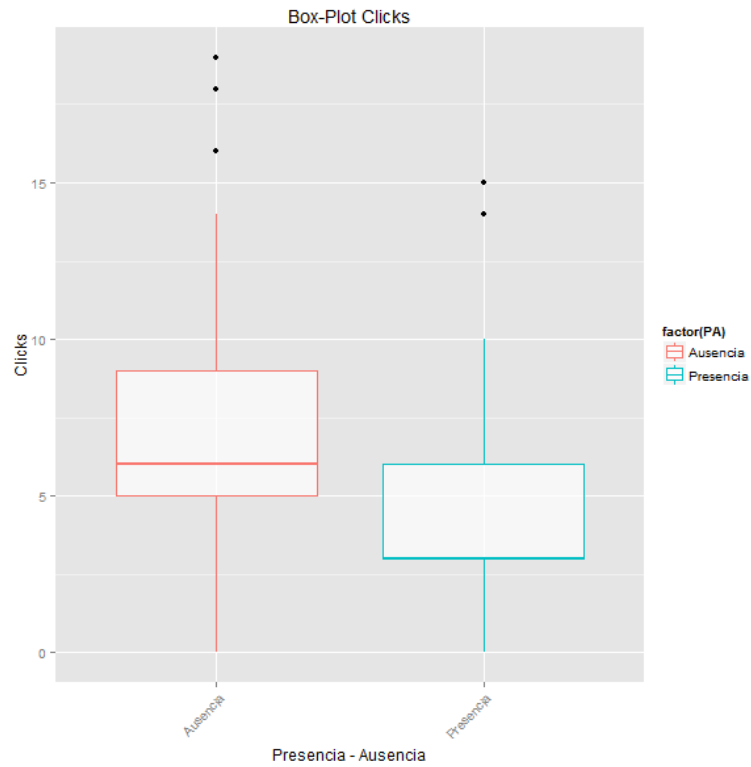


Figura 6: Diagrama de cajas para el Número de Clicks en el mecanismo Retroalimentación de Progreso

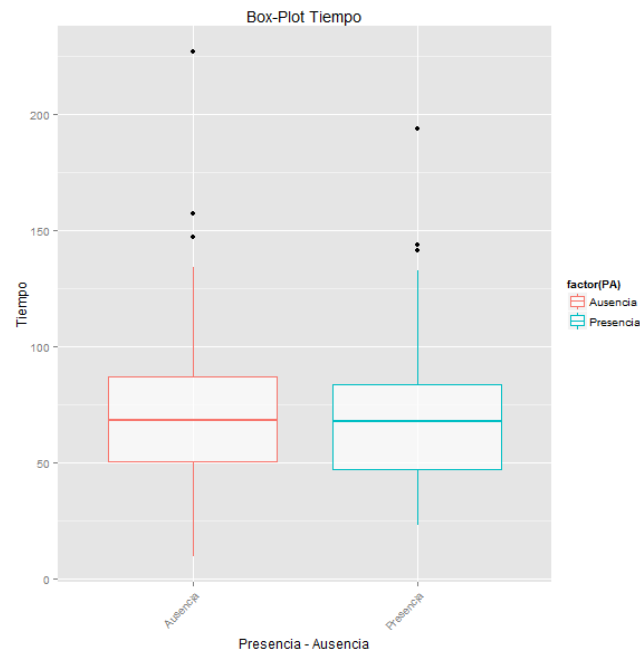


Figura 7: Diagrama de cajas para el Tiempo en el mecanismo Retroalimentación de Progreso

3.11.1.3. Abortar Operación

En la Tabla 17, en la que se muestran los estadísticos descriptivos para la variable respuesta eficiencia sobre el mecanismo de usabilidad Abortar Operación, se observa que, en el caso del número de clicks la media es menor en el caso de presencia del mecanismo frente al caso de ausencia del mismo. La desviación típica en ambos casos es muy similar. En cuanto a los coeficientes de asimetría y Curtosis todos son positivos.

Fijándonos en el diagrama de caja para esta variable, Figura 8, podemos observar como la muestra que no dispone del mecanismo tiene tres valores atípicos frente a la muestra de los sujetos que sí disponen del mecanismo, que tiene el doble de *outliers*.

Si analizamos la variable tiempo, comprobamos, en este caso, que la media es muy superior en el caso de la ausencia del mecanismo, en comparación con la media para aquellos sujetos que sí disponían del mecanismo Abortar Operación (Tabla 17). En este caso, la desviación típica también es muy superior cuando el mecanismo no está disponible.

Sin embargo, el estadístico que destaca más es el mínimo: el sujeto que menos tiempo tardó en hacer la tarea y que además disponía del MU-ABO lo hizo en 5.34 segundos en contraposición al usuario que menos segundos necesitó para completar la tarea, sin el mecanismo, que fue en 43.25 segundos, es decir, casi ocho veces más.

En cuanto al diagrama de cajas de la Figura 9, se observa que en el caso de presencia del mecanismo hay dos valores atípicos y en la población que no disponía del mismo, hay el doble, cuatro.

Es importante aclarar que en ambos casos, tanto para el tiempo como para el número de clicks, en la muestra se encuentran valores imposibles: por ejemplo, algún sujeto de la población Clicks – Presencia dice haber realizado la tarea con cero clicks, lo que es completamente imposible. Todos estos valores serán eliminados antes de realizar el test ANOVA, tal y como se ha comentado anteriormente.

| ABO | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|----------|-------|-------|--------|---------|-------|-------|--------|--------|-------|--------|----------|----------|----------|
| Clicks P | 0 | 81 | 16.79 | 11 | 11 | 7 | 11 | 18 | 81 | 16.86 | 284.08 | 2.23 | 5.25 |
| Clicks A | 1 | 78 | 20.41 | 15 | 18.5 | 9 | 15 | 27.5 | 78 | 16.76 | 280.92 | 1.53 | 2.31 |
| Tiempo P | 5.34 | 447.4 | 124.75 | 103.21 | 92.17 | 64.67 | 103.21 | 156.83 | 447.4 | 83.26 | 6933.31 | 1.78 | 4.38 |
| Tiempo A | 43.25 | 545 | 160.01 | 141.07 | 94.26 | 90.51 | 141.07 | 184.77 | 545 | 100.00 | 10000.61 | 1.85 | 4.13 |

Tabla 17: Estadísticos descriptivos para el mecanismo Abortar Operación en el estudio de Eficiencia

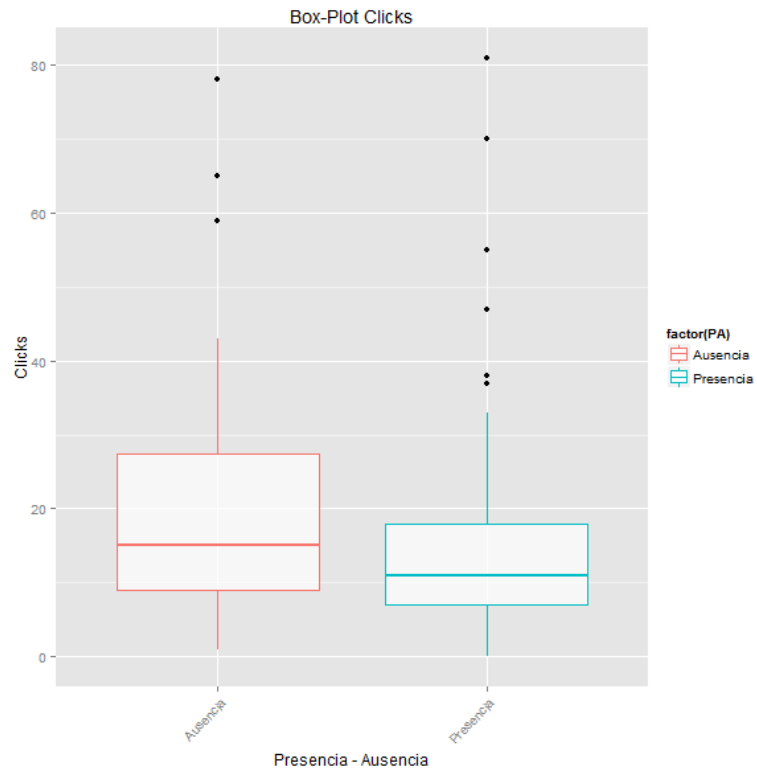


Figura 8: Diagrama de cajas para el Número de Clicks en el mecanismo Abortar Operación

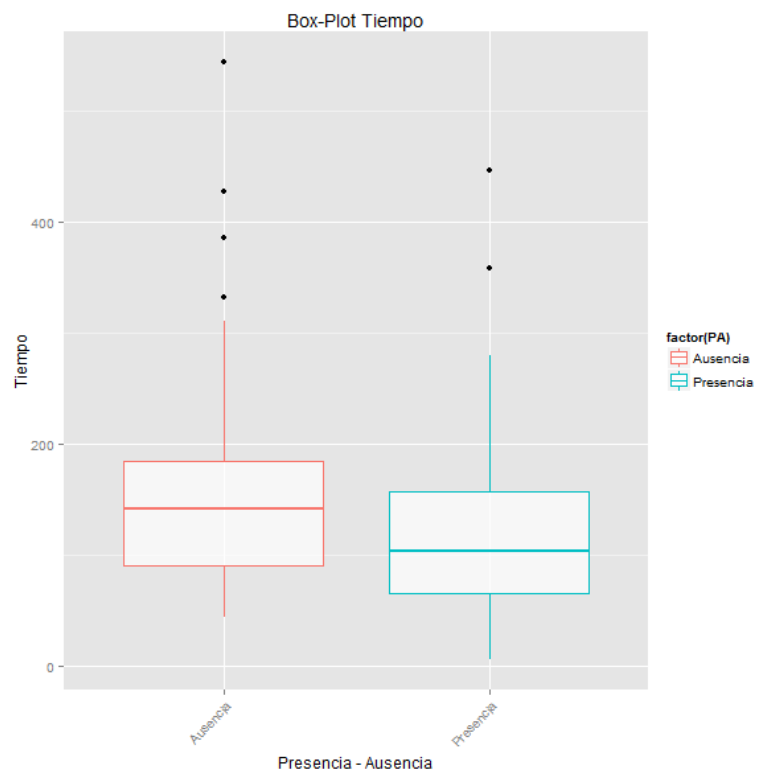


Figura 9: Diagrama de cajas para el Tiempo en el mecanismo Abortar Operación

3.11.2. Satisfacción

3.11.2.1. Preferencias

En la Tabla 18, se observa como la media en el caso de ausencia del mecanismo es más de un punto inferior que en el caso de la presencia del MU-PRE. De la misma forma, la mediana es superior en el caso de la presencia del mecanismo y lo mismo sucede con cualquiera de los demás cuartiles. Fijándonos en la desviación típica, se observa que en ambos casos es algo superior a un punto. En cuanto a los coeficientes de asimetría y Curtois, en el caso en el que el mecanismo esté activo se obtienen coeficientes negativos, mientras que en el caso en el que no esté disponible el mecanismo, los coeficientes son positivos.

| PRE | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|---------|------|------|-------|---------|------|----|------|------|------|------|----------|----------|----------|
| Valor P | 1 | 5 | 3.36 | 3.75 | 2.5 | 2 | 3.75 | 4.5 | 5 | 1.35 | 1.82 | -0.38 | -1.23 |
| Valor A | 1 | 4.75 | 1.77 | 1.25 | 1.06 | 1 | 1.25 | 2.06 | 4.75 | 1.08 | 1.16 | 1.37 | 0.75 |

Tabla 18: Estadísticos descriptivos para el mecanismo Preferencias para el estudio de Satisfacción

En el diagrama de cajas representado en la Figura 10, se observa cómo hay bastante diferencia entre ambas poblaciones, aunque ninguna de ellas sea simétrica. Es destacable que en el caso de la ausencia del mecanismo hay tres valores atípicos, mientras que en el caso de la presencia del mecanismo no se obtiene ninguno.

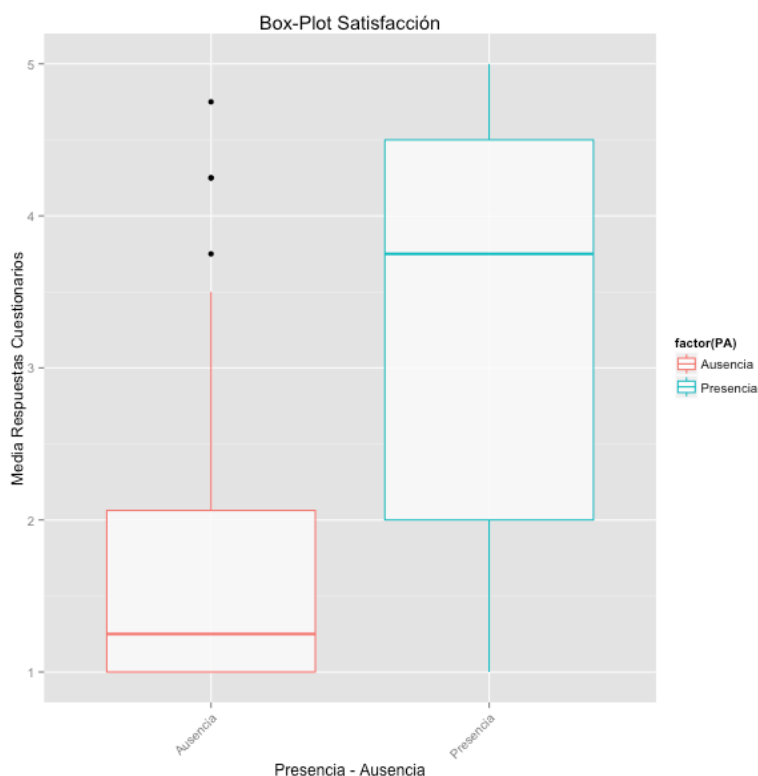


Figura 10: Diagrama de cajas para la Satisfacción en el mecanismo Preferencias

3.11.2.2. Retroalimentación de Progreso

En el caso del mecanismo de usabilidad Retroalimentación de Progreso, si se estudian los estadísticos descriptivos, se observa un comportamiento muy similar en las dos poblaciones, Tabla 19. Lo más destacable es que el coeficiente de asimetría es negativo en el caso de la presencia del mecanismo mientras que, es positivo, aunque muy próximo a cero en la población en la que el mecanismo está ausente. En el coeficiente de Curtois el comportamiento es el contrario: cuando el mecanismo está presente es positivo, mientras que cuando el mecanismo está ausente es negativo.

| PRO | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|---------|------|------|-------|---------|------|------|------|------|----|------|----------|----------|----------|
| Valor P | 1.33 | 5 | 4.30 | 4.33 | 1 | 4 | 4.33 | 5 | 5 | 0.68 | 0.46 | -1.74 | 5.70 |
| Valor A | 3 | 5 | 4.02 | 4 | 0.66 | 3.66 | 4 | 4.33 | 5 | 0.55 | 0.31 | 0.24 | -0.88 |

Tabla 19: Estadísticos descriptivos para el mecanismo Retroalimentación de Progreso para el estudio de Satisfacción

En la Figura 11 se puede comprobar como ambas poblaciones están centradas en los valores superiores del cuestionario. Sin embargo, la población para la ausencia del mecanismo es bastante simétrica, pero no ocurre lo mismo para la población con presencia del mecanismo. Además, en el caso de presencia del mecanismo, no hay bigote superior y el bigote inferior de ambas poblaciones termina en el mismo punto, 3.

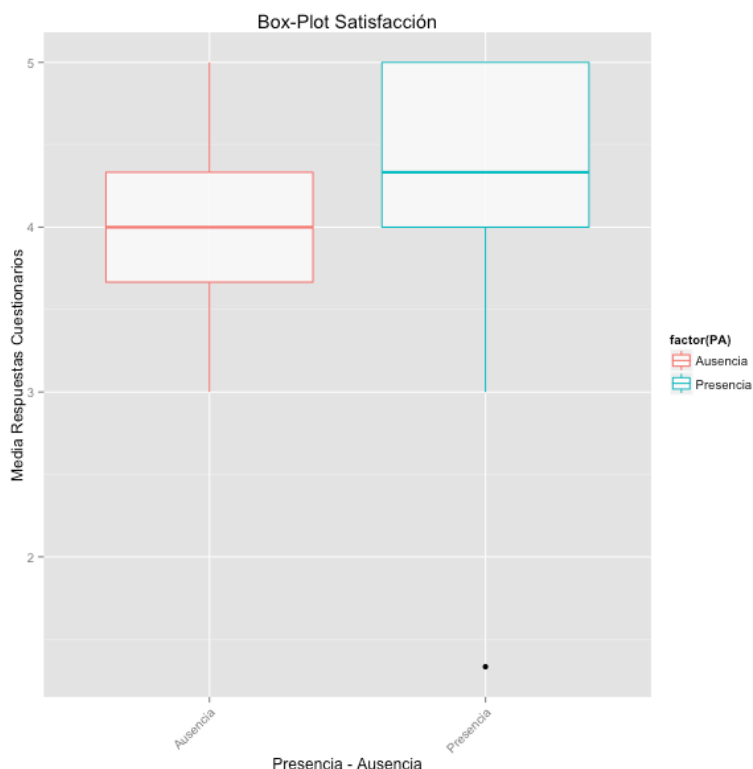


Figura 11: Diagrama de cajas para la Satisfacción en el mecanismo Retroalimentación de Progreso

3.11.2.3. Abortar Operación

Finalmente, analicemos qué sucede con el mecanismo Abortar Operación para terminar de estudiar los estadísticos descriptivos de la variable satisfacción. Fijándonos en la Tabla 20, podemos comprobar como el mínimo y el máximo coinciden en ambos casos. La media ya es distinta siendo superior en la población de los sujetos que disponen del mecanismo ABO. La mediana supera en dos puntos en Presencia frente a Ausencia. Sin embargo, la desviación típica es muy similar en los dos casos. Cabe destacar también que el coeficiente de asimetría en el caso de presencia del mecanismo es negativo, mientras que, en ausencia del mecanismo es positivo. El coeficiente de Curtosis es negativo en ambos casos.

| ABO | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|---------|------|------|-------|---------|------|------|----|------|----|------|----------|----------|----------|
| Valor P | 1 | 5 | 3.69 | 4 | 1.83 | 2.83 | 4 | 4.66 | 5 | 1.30 | 1.70 | -0.84 | -0.58 |
| Valor A | 1 | 5 | 2.31 | 2 | 2 | 1.33 | 2 | 3.33 | 5 | 1.21 | 1.47 | 0.71 | -0.62 |

Tabla 20: Estadísticos descriptivos para el mecanismo Abortar Operación para el estudio de Satisfacción

El diagrama de cajas, Figura 12, nos ayuda a terminar de comprender la forma de las dos muestras. Se podría decir que una muestra es la inversa de la otra, en el sentido de que en ausencia, el bigote inferior es pequeño, en presencia el bigote superior es pequeño. En ausencia, el bigote superior es más alargado, en presencia, el bigote inferior es más alargado. En ausencia la distancia del Q1 al Q2 o mediana es menor que la distancia de Q2 a Q3, justo lo contrario que sucede en presencia del mecanismo. En ningún caso se presentan valores atípicos.

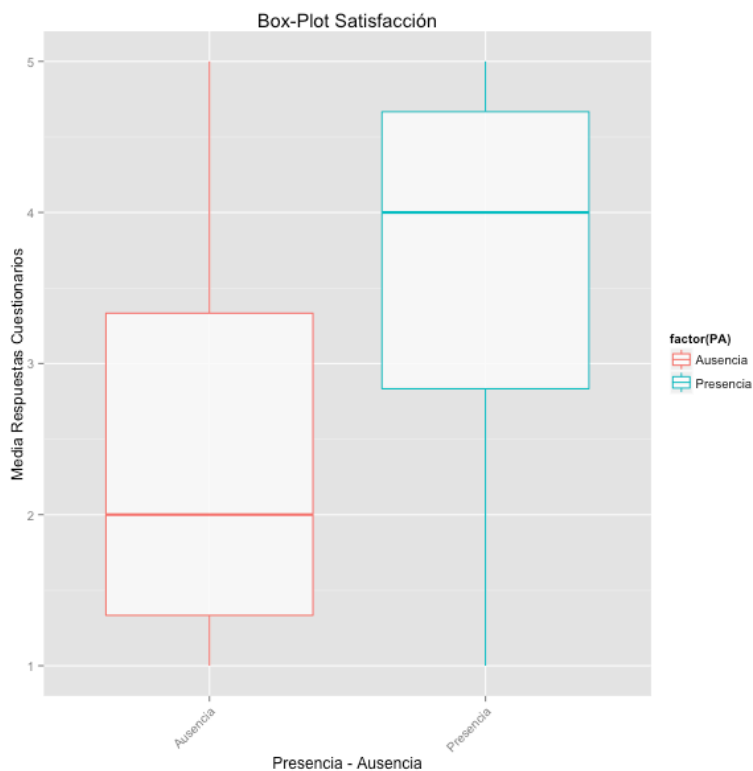


Figura 12: Diagrama de cajas para la Satisfacción en el mecanismo Abortar Operación

3.12. ANOVA

En esta sección se presentan los resultados obtenidos para los test ANOVA realizados después de analizar los estadísticos descriptivos. Como se ha comentado en la sección 3.10 para realizar el test ANOVA es necesario comprobar previamente las hipótesis del mismo. Éstas son:

1. Las observaciones deben ser independientes.
2. Las poblaciones deben ser normales.
3. Las poblaciones deben ser homocedásticas, es decir, todas deben tener igual varianza.

La primera de ellas se obtiene por construcción del experimento. Todos los sujetos han realizado sus tareas en ordenadores individuales y sin contacto con el resto de los participantes. Además, las tareas de unos no estaban influenciadas por las tareas de los demás.

En cuanto a la segunda y tercera hipótesis será necesario comprobarlas en cada caso y, por tanto, se incluirán los resultados al inicio de la subsección pertinente. Para comprobar la normalidad de residuos se realizará el test de Shapiro-Wilk, mientras que, para comprobar la homocedasticidad se utilizará el test de Levene.

Por último, recordar el tipo de test ANOVA que se ha realizado. Se trata de un test ANOVA Factorial Fraccionado conforme a la Tabla 14. En él se tienen en cuenta los factores Presencia/Ausencia, Orden y Grupo. Para la realización de todo el estudio estadístico se ha empleado el lenguaje R [Field et al., 2012], que permite seleccionar los niveles de significación deseados. En la Tabla 21 se presentan los niveles de significación y símbolos de reconocimiento tenidos en cuenta.

| Nivel de Significación | Símbolo de Reconocimiento |
|------------------------|---------------------------|
| $0 \leq p < 0.001$ | *** |
| $0.001 \leq p < 0.01$ | ** |
| $0.01 \leq p < 0.05$ | * |

Tabla 21: Distintos niveles de significación considerados

3.12.1. Eficiencia

3.12.1.1. Preferencias

Comenzamos el análisis de la eficiencia para el mecanismo Preferencias utilizando la variable tiempo. Es necesario comentar que las poblaciones han necesitado ser transformadas para poder satisfacer las hipótesis del test ANOVA [Field et al., 2012]. En este caso se ha utilizado la $X \rightarrow \log(X)$. Como se desprende de la Tabla 22, las poblaciones, bajo esta transformación, satisfacen las hipótesis de normalidad ($p\text{-valor} = 0.3541 > 0.05$) y homogeneidad ($p\text{-valor} = 0.8394 > 0.05$). En la Tabla 23 se observa cómo la Presencia o Ausencia del mecanismo de usabilidad no arroja información concluyente. El test ANOVA determina que el factor más significativo es el Orden de exposición pues es el factor que obtiene un valor menor para el $p\text{-valor}$.

Para comprobar los distintos comportamientos del Orden se utiliza el test de Tukey [Juristo & Moreno, 2001]. En la Figura 40, contenida en el anexo E, se puede comprobar como existen distintos grupos de comportamiento. El intervalo Orden5–Orden1 tiene un *p-valor* de 0.0033 y, también, el intervalo Orden5–Orden2 tiene un *p-valor* de 0.0247. Esto supone, tal y como se evidenciaba en el test ANOVA, que hay dos grupos en órdenes que tienen comportamientos diferenciados: {1, 2} y {5}.

Estudiando qué sucede en la variable número de clicks observamos que el resultados es bastante similar. En primer lugar, nuevamente la población inicial no verifica las hipótesis pero con la misma transformación que en el caso anterior, $X \rightarrow \log(X)$, no hay ningún problema para pasar los test de normalidad y homogeneidad tal y como se puede observar en Tabla 24. En la Tabla 25 se comprueba cómo el único factor concluyente en el test ANOVA es el Orden. Si se plantea el test de Tukey se obtiene el gráfico de la Figura 41, en el que se observa que los órdenes 5 y 2 tienen comportamientos diferentes, concretamente el *p-valor* para este intervalo es 0.0311.

En conclusión, se debe aceptar la hipótesis H.1.3.0, es decir, bajo este experimento se determina que no existe diferencia significativa en la eficiencia del usuario al incorporar o no el mecanismo de usabilidad Preferencias.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9784 | 0.3541 | 0.6744 | 0.8394 |

Tabla 22: Test de normalidad y homocedasticidad variable Tiempo MU-PRE

| Tiempo | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|-------------|----|--------|---------|---------|---------|----|
| PA | 1 | 0.111 | 0.1105 | 0.679 | 0.42504 | |
| Grupo | 2 | 0.078 | 0.0389 | 0.239 | 0.78843 | |
| Orden | 5 | 3.585 | 0.7169 | 4.407 | 0.00301 | ** |
| PA:Orden | 5 | 0.240 | 0.0480 | 0.295 | 0.91251 | |
| Grupo:Orden | 10 | 2.919 | 0.2929 | 1.800 | 0.09500 | . |
| Residuos | 37 | 6.018 | 0.1627 | | | |

Tabla 23: Test de ANOVA para la variable Tiempo en MU-PRE

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9852 | 0.402 | 0.9222 | 0.5707 |

Tabla 24: Test de normalidad y homocedasticidad variable N° de Clicks MU-PRE

| Clicks | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|-------------|----|--------|---------|---------|---------|---|
| PA | 1 | 0.18 | 0.185 | 0.160 | 0.6903 | |
| Grupo | 2 | 0.41 | 0.206 | 0.179 | 0.8366 | |
| Orden | 5 | 16.20 | 3.239 | 2.809 | 0.0232 | * |
| PA:Orden | 5 | 0.91 | 0.182 | 0.158 | 0.9769 | |
| Grupo:Orden | 10 | 1.84 | 0.184 | 0.160 | 0.9983 | |
| Residuos | 66 | 76.10 | 1.153 | | | |

Tabla 25: Test de ANOVA para la variable N° de Clicks en MU-PRE

3.12.1.2. Retroalimentación de Progreso

En esta sección se presentan los resultados para la eficiencia y el mecanismo de usabilidad Retroalimentación de Progreso. En primer lugar, resaltar que tanto en el caso de la variable tiempo como en la variable número de clicks ha sido necesario realizar una transformación de las poblaciones para conseguir superar las hipótesis del test ANOVA. Para el tiempo se ha aplicado la transformación $X \rightarrow \log(X)$ mientras que para la variable número de clicks se ha aplicado la transformación $X \rightarrow 1/X$.

Con estas transformaciones, la variable tiempo pasa las hipótesis de normalidad y homogeneidad sin dificultades, tal y como se muestra en la Tabla 26. Sin embargo, la Tabla 28, muestra como, a pesar de realizar la transformación mencionada, en el caso de la variable número de clicks, la normalidad no se consigue aunque el *p-valor* se queda cerca de 0.05. No es un problema sumamente importante, pues el *p-valor* para el test de Levene es 0.8798 lo que determina que supera el test holgadamente y, además, el test ANOVA no es sensible a la normalidad bajo esta situación de homogeneidad.

Estudiando ahora el test ANOVA para la variable tiempo, Tabla 27, se observa que el único factor determinante es el Orden de exposición. Tras este resultado, se realiza el test de Tukey para entender en qué órdenes existen comportamientos notablemente diferentes. Tal y como se muestra en la Figura 42, aunque el test ANOVA refleja que el factor Orden es significativo, el test de Tukey muestra que las diferencias entre los grupos son mínimas pues, en todos los intervalos se encuentra el valor cero.

Por último, analizando el test ANOVA para la variable número de clicks, Tabla 29, se obtiene que hay dos factores significativos la Presencia/Ausencia y el Orden de exposición. Resulta necesario verificar qué sucede para el Orden y, como en otras situaciones, se ha utilizado el test de Tukey. En la Figura 43 se muestra la gráfica de este test de la que se desprende que hay diferencias significativas entre el intervalo Orden 4-Orden 6, *p-valor* = 0.0176, y el intervalo Orden1-Orden6, *p-valor* = 0.03110.

Se puede concluir, por tanto, un comportamiento diferente en las dos formas de medir la eficiencia. En el primer caso, tiempo de realización de la tarea, se acepta la hipótesis nula H.1.2.0 que indica que no hay diferencia significativa entre la presencia o ausencia del mecanismo. Sin embargo, en el segundo caso, nivel de interacción del sistema, se rechaza la hipótesis H.1.2.0, aceptando la hipótesis alternativa H.1.2.1 y, por tanto, se

determina que existe diferencia significativa en la eficiencia del usuario al incorporar, o no, el mecanismo de usabilidad Retroalimentación de Progreso.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9909 | 0.7583 | 1.0556 | 0.4138 |

Tabla 26: Test de normalidad y homocedasticidad variable Tiempo MU-PRO

| Tiempo | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|-------------|----|--------|---------|---------|---------|---|
| PA | 1 | 0.092 | 0.0919 | 0.498 | 0.4827 | |
| Grupo | 2 | 0.640 | 0.3201 | 1.734 | 0.1839 | |
| Orden | 5 | 2.187 | 0.4374 | 2.370 | 0.0477 | * |
| PA:Orden | 5 | 0.413 | 0.0827 | 0.448 | 0.8134 | |
| Grupo:Orden | 10 | 1.754 | 0.1754 | 0.950 | 0.4938 | |
| Residuos | 72 | 13.290 | 0.1846 | | | |

Tabla 27: Test de ANOVA para la variable Tiempo en MU-PRO

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9696 | 0.02721 | 0.6455 | 0.8798 |

Tabla 28: Test de normalidad y homocedasticidad variable N° de Clicks MU-PRO

| Clicks | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|-------------|----|--------|---------|---------|----------|-----|
| PA | 1 | 0.1412 | 0.14124 | 22.915 | 9.13e-06 | *** |
| Grupo | 2 | 0.0058 | 0.00292 | 0.473 | 0.62492 | |
| Orden | 5 | 0.1122 | 0.02244 | 3.641 | 0.00551 | ** |
| PA:Orden | 5 | 0.0502 | 0.01003 | 1.627 | 0.16412 | |
| Grupo:Orden | 10 | 0.0523 | 0.00523 | 0.848 | 0.58500 | |
| Residuos | 70 | 0.4315 | 0.00616 | | | |

Tabla 29: Test de ANOVA para la variable N° de Clicks en MU-PRO

3.12.1.3. Abortar Operación

Para la variable tiempo, después de realizar la transformación $X \rightarrow \log(X)$, no existe ningún problema a la hora de comprobar la normalidad de residuos y la homogeneidad de varianzas de la misma, Tabla 30. Sin embargo, en el test ANOVA no se obtiene el mejor resultado posible. La Tabla 31 muestra cómo se obtienen resultados óptimos tanto para el factor Presencia/Ausencia como para el factor Orden. Debemos, por tanto, estudiar qué sucede con el factor Orden tal y como se ha hecho en situaciones anteriores. Utilizando el test de Tukey, representado en la Figura 44 dentro del anexo E,

se observa que todos los intervalos contienen el 0, lo que supone que en cierto momento todas las muestras pueden tomar valores iguales. Lo que hace suponer que, aunque haya pasado el test ANOVA la variable Orden, ésta no tiene mucha influencia.

Del test ANOVA también se desprende que el factor Presencia/Ausencia es determinante.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9856 | 0.3807 | 0.4715 | 0.9772 |

Tabla 30: Test de normalidad y homocedasticidad variable Tiempo MU-ABO

| Tiempo | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|-------------|----|--------|---------|---------|---------|----|
| PA | 1 | 2.250 | 2.2495 | 8.238 | 0.00538 | ** |
| Grupo | 2 | 0.168 | 0.0842 | 0.308 | 0.73569 | |
| Orden | 5 | 4.460 | 0.8920 | 3.266 | 0.01029 | * |
| PA:Orden | 5 | 0.884 | 0.1768 | 0.647 | 0.66442 | |
| Grupo:Orden | 10 | 2.295 | 0.2295 | 0.840 | 0.49181 | |
| Residuos | 72 | 19.662 | 0.2731 | | | |

Tabla 31: Test de ANOVA para la variable Tiempo en MU-ABO

Analizamos ahora el caso de la variable número de clicks. Aquí nuevamente se superan las hipótesis de normalidad y homogeneidad sin problemas, tal y como se observa en la Tabla 32, después de realizar la transformación $X \rightarrow 1/X$. Fijándose en la Tabla 33, se puede comprobar como en el test ANOVA el único factor decisivo es la Presencia/Ausencia del mecanismo Abortar Operación, pues el *p-valor* es $0.0222 < 0.05$.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9838 | 0.3682 | 0.7331 | 0.7931 |

Tabla 32: Test de normalidad y homocedasticidad variable N° de Clicks MU-ABO

| Clicks | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|-------------|----|---------|----------|---------|---------|---|
| PA | 1 | 0.00852 | 0.008519 | 5.508 | 0.0222 | * |
| Grupo | 2 | 0.00105 | 0.000524 | 0.339 | 0.7139 | |
| Orden | 5 | 0.00644 | 0.001289 | 0.833 | 0.5313 | |
| PA:Orden | 5 | 0.00306 | 0.000611 | 0.395 | 0.8503 | |
| Grupo:Orden | 10 | 0.02313 | 0.002313 | 1.496 | 0.1631 | |
| Residuos | 61 | 0.09434 | 0.001547 | | | |

Tabla 33: Test de ANOVA para la variable N° de Clicks en MU-ABO

En conclusión, existen diferencias significativas en la eficiencia del usuario al incorporar o no el mecanismo de usabilidad Abortar Operación. Es decir, se debe rechazar la hipótesis nula H.1.1.0 y aceptar la hipótesis alternativa H.1.1.1 de la Tabla 8.

3.12.2. Satisfacción

3.12.2.1. Preferencias

En la Tabla 34, se observa como se supera tanto la hipótesis de normalidad como la de homocedasticidad para la media de respuestas del cuestionario post-tarea para el mecanismo Preferencias. Además, en la Tabla 35, se determina como tras realizar el test ANOVA se debe rechazar la hipótesis nula planteada en el experimento, concretamente la Hipótesis H.2.3.0 reflejada en la Tabla 8. Esto supone que se acepta la hipótesis H.2.3.1, que determina que existen diferencias significativas en la satisfacción del usuario al incorporar el mecanismo Preferencias o al no incorporarlo. Se puede decir, además, que las diferencias son grandes pues el *p-valor*, para esta variable, toma un número muy pequeño, 1.21e-08.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9747 | 0.05124 | 1.2144 | 0.2599 |

Tabla 34: Test de normalidad y homocedasticidad para la Satisfacción en MU-PRE

| Satisfacción | DF | Sum Sq | Mean Sq | F Value | Pr (>F) |
|--------------|----|--------|---------|---------|--------------|
| PA | 1 | 62.11 | 62.11 | 39.194 | 1.21e-08 *** |
| Grupo | 2 | 1.30 | 0.65 | 0.411 | 0.664 |
| Orden | 1 | 0.19 | 0.19 | 0.123 | 0.727 |
| PA:Orden | 1 | 2.63 | 2.63 | 1.659 | 0.201 |
| Grupo:Orden | 2 | 0.36 | 0.18 | 0.113 | 0.893 |
| Residuos | 92 | 145.79 | 1.58 | | |

Tabla 35: Test de ANOVA para la Satisfacción en MU-PRE

3.12.2.2. Retroalimentación de Progreso

Nuevamente, como en el estudio de la eficiencia, aquí ha sido necesario realizar una transformación de los datos, de acuerdo a [Field et al., 2012]. La transformación concreta ha sido $X \rightarrow \log(X)$.

Observando los resultados obtenidos, en primer lugar, hay que destacar, tal y como refleja la Tabla 36, que la hipótesis de normalidad casi se consigue pues toma un valor muy próximo a 0.05. Se considera válida esta hipótesis pues muchos autores consideran que, para el test ANOVA, es suficiente con superar 0.01, pues no es un test muy comprometido con la normalidad, siempre que la homocedasticidad esté completamente

asegurada. Esto último sucede, con creces, en nuestro caso, pues el *p-valor* en el test de Levene alcanza 0.9079.

Fijándonos ahora en la Tabla 37, se observa como al igual que en el caso del mecanismo Preferencias, en este caso se rechaza la hipótesis H.2.2.0 en favor de la hipótesis H.2.2.1, de la Tabla 8, con lo que se concluye que existe diferencia significativa en la satisfacción del usuario entre incorporar el mecanismo de usabilidad Retroalimentación de Progreso y no incorporarlo.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9726 | 0.03653 | 0.611 | 0.9079 |

Tabla 36: Test de normalidad y homocedasticidad para la Satisfacción en MU-PRO

| Satisfacción | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|--------------|----|--------|---------|---------|---------|----|
| PA | 1 | 0.1765 | 0.17654 | 9.393 | 0.00287 | ** |
| Grupo | 2 | 0.0277 | 0.01387 | 0.738 | 0.48090 | |
| Orden | 1 | 0.0281 | 0.02813 | 1.497 | 0.22436 | |
| PA:Orden | 1 | 0.0000 | 0.00004 | 0.002 | 0.96164 | |
| Grupo:Orden | 2 | 0.0249 | 0.01245 | 0.662 | 0.51817 | |
| Residuos | 91 | 1.7104 | 0.01880 | | | |

Tabla 37: Test de ANOVA para la Satisfacción en MU-PRO

3.12.2.3. Abortar Operación

Para completar el estudio de la variable respuesta satisfacción, nos fijamos ahora en el mecanismo de usabilidad Abortar Operación.

Este resultado es muy similar al obtenido cuando se analizó el mecanismo Preferencias. Sin realizar ningún tipo de transformación a los datos, se supera fácilmente el test de normalidad y también el test de homocedasticidad, Tabla 38.

En cuanto al test ANOVA se observa como la variable determinante es la Presencia/Ausencia del mecanismo, Tabla 39. Por tanto, se concluye que es necesario rechazar la hipótesis nula H.2.1.0, es decir, aceptar la hipótesis alternativa H.2.1.1, ambas expuestas en la Tabla 8, con lo que se determina que existe diferencia significativa en la satisfacción del usuario entre incorporar el mecanismo Abortar Operación y no incorporarlo.

| Test Shapiro-Wilk | | Test de Levene | |
|-------------------|---------|----------------|---------|
| Estadístico | P-valor | Estadístico | P-valor |
| 0.9777 | 0.09107 | 0.4761 | 0.9735 |

Tabla 38: Test de normalidad y homocedasticidad para la Satisfacción en MU-ABO

| Satisfacción | DF | Sum Sq | Mean Sq | F Value | Pr (>F) | |
|---------------------|-----------|---------------|----------------|----------------|-------------------|-----|
| <i>PA</i> | 1 | 47.14 | 47.14 | 30.009 | 3.79e-07 | *** |
| <i>Grupo</i> | 2 | 3.15 | 1.57 | 1.003 | 0.371 | |
| <i>Orden</i> | 1 | 3.29 | 3.29 | 2.096 | 0.151 | |
| <i>PA:Orden</i> | 1 | 1.56 | 1.56 | 0.996 | 0.321 | |
| <i>Grupo:Orden</i> | 2 | 3.10 | 1.55 | 0.987 | 0.377 | |
| <i>Residuos</i> | 91 | 142.95 | 1.57 | | | |

Tabla 39: Test de ANOVA para la Satisfacción en MU-ABO

CAPÍTULO 4. DISCUSIÓN DE RESULTADOS

En este capítulo se realiza la discusión de los resultados obtenidos en esta investigación. Por esta razón, en este punto se comentarán los resultados del test ANOVA en relación con las hipótesis que se aceptan.

4.1. Discusión

Tras analizar los datos obtenidos en los diferentes test, presentados en la sección 3.12, se puede realizar un resumen básico de los resultados, contenido en la Tabla 40, que ayude a comprender las conclusiones de la investigación.

| Variable Respuesta | | PRE | | PRO | | ABO | |
|--------------------|--------|-----------|--------|-----------|-----------------|-----------|---------------|
| | | Hipótesis | ANOVA | Hipótesis | ANOVA | Hipótesis | ANOVA |
| Eficiencia | Tiempo | H.1.3.0 | OR ** | H.1.2.0 | OR * | H.1.1.1 | PA ** OR * |
| | Clicks | H.1.3.0 | OR * | H.1.2.1 | PA *** OR ** | H.1.1.1 | PA* |
| Satisfacción | | H.2.3.1 | PA *** | H.2.2.1 | PA ** | H.2.1.1 | PA *** |

Tabla 40: Resumen de resultados del experimento con ANOVA

Es decir, a nivel de hipótesis de investigación, se concluye:

1. Se acepta la hipótesis nula H.1.3.0 que determina que no existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRE o al no incorporarlo.
2. Esta situación es la más controvertida. Se acepta la hipótesis alternativa H.1.2.1 que determina que existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRO o al no incorporarlo, cuando se calcula la eficiencia utilizando la variable número de clicks. Sin embargo, se debe aceptar la hipótesis nula H.1.2.0 que determina que no existe diferencia significativa en la eficiencia del usuario al incorporar el MU-PRO o al no incorporarlo, cuando se calcula la eficiencia utilizando la variable tiempo.
3. Se acepta la hipótesis alternativa H.1.1.1 que determina que existe diferencia significativa en la eficiencia del usuario al incorporar el MU-ABO o al no incorporarlo.
4. Se acepta la hipótesis alternativa H.2.3.1 que determina que existe diferencia significativa en la satisfacción del usuario al incorporar el MU-PRE o al no incorporarlo.

5. Se acepta la hipótesis alternativa H.2.2.1 que determina que existe diferencia significativa en la satisfacción del usuario al incorporar el MU-PRO o al no incorporarlo.
6. Se acepta la hipótesis alternativa H.2.1.1 que determina que existe diferencia significativa en la satisfacción del usuario al incorporar el MU-ABO o al no incorporarlo.

Es necesario profundizar algo más en los casos en los que no se puede aceptar la hipótesis alternativa.

Tanto en el caso de la variable tiempo como en el caso de la variable número de clicks para el mecanismo Preferencias el test ANOVA determina que el factor Orden es el determinante (ver Tabla 25 y Tabla 23 respectivamente).

Si analizamos la variable tiempo para el mecanismo Retroalimentación de Progreso, el factor significativo es el Orden y, en el caso de la variable número de clicks existen dos factores determinantes, Presencia/Ausencia y Orden (siendo menor el *p-valor* en el caso de Presencia/Ausencia). Sin embargo, como se muestra en la Figura 42, el test de Tukey para el factor Orden y la variable tiempo determina que realmente el orden no tiene un comportamiento diferente. Si nos fijásemos en el test de Tukey para la variable clicks, aquí sí que se observa una diferencia aunque mínima, Figura 43. Es por ello, que realmente sería más correcto concluir que en el caso de la variable número de clicks Presencia/Ausencia es el factor más determinante y, en el caso de la variable tiempo no se observan diferencias significativas altas.

Finalmente, mencionar que en el caso de la variable tiempo para el mecanismo Abortar Operación, tal y como se muestra en la Figura 44 y en la Tabla 31 el orden parece ser significativo durante el test ANOVA pero después con el test de Tukey se comprueba que esta significación es casi nula y, por esta razón, se acepta la hipótesis alternativa H.1.1.1 que se ha presentado anteriormente.

4.2. Interpretación

Llegado a este punto se comentará la interpretación de los resultados discutidos en la sección precedente. Esta interpretación se organizará por variables respuesta analizadas.

De acuerdo a la Tabla 40, incorporar los mecanismos de usabilidad Preferencias, Retroalimentación de Progreso y Abortar Operación siempre influye positivamente en la satisfacción de los usuarios finales. Éste es un resultado intuitivamente lógico pero con un carácter importante pues de este experimento se desprende una razón de peso a favor de la implementación de dichos mecanismos.

De la misma forma, la eficiencia de los usuarios mejora cuando el mecanismo Abortar Operación está presente tanto en el caso de la eficiencia vista como la rapidez con la que se realiza una tarea, como la eficiencia entendida como el nivel de interacción con el sistema. Esta situación vuelve a ser lógica y se incorpora a la anterior para concluir que las mejoras obtenidas al proveer a una aplicación web con alto contenido interactivo del

mecanismo Abortar Operación son percibidas tanto a nivel de satisfacción, rapidez y nivel de interacción con la herramienta.

La eficiencia medida como el nivel de interacción con el sistema también mejora cuando los usuarios disponen del mecanismo Retroalimentación de Progreso. Sin embargo, no se han encontrado evidencias significativas que hagan pensar que la rapidez mejora cuando se dispone de este mecanismo. Estos resultados hacen que directamente debamos hacernos la siguiente pregunta:

¿Por qué el factor Orden es determinante para la rapidez del usuario en el MU-PRO?

Como se ha comentado en la sección precedente, en este caso el test ANOVA puede llevar a engaño pues, aunque parece que el factor Orden es significativo y concluyente, en el test de Tukey se determina que realmente no se pueden identificar comportamientos significativamente diferentes entre los distintos órdenes de exposición.

Por tanto, se debe remarcar que el problema real en este caso es que los datos recogidos durante el experimento no resultan determinantes para ninguno de los factores estudiados. Esto se puede deber a que tanto si el mecanismo está presente como ausente, el tiempo de espera del usuario mientras se realiza la consulta de búsqueda del artículo es el mismo. Además, el resto de la tarea es muy sencilla y, por tanto, en ambos casos el tiempo utilizado puede tender a ser similar, sin ser un resultado ilógico. Es importante destacar que aunque el tiempo es el mismo, sí se observa que los usuarios que no disponen del mecanismo Retroalimentación de Progreso hacen muchos más clicks que se pueden entender que en un cierto momento, el momento de la consulta, pierden el control de qué sucede e incluso piensan que la herramienta ha dejado de funcionar correctamente.

Por último, la situación, más crítica es la relativa a la eficiencia para el mecanismo Preferencias. Para ninguna de las dos variables estudiadas la Presencia o Ausencia del mecanismo resulta determinante.

¿Por qué ambos test ANOVA concluyen que el factor diferenciador es el Orden?

En primer lugar analicemos la rapidez. En este caso el test de Tukey refleja que hay órdenes con comportamientos diferentes y esto se puede deber al aprendizaje que van adquiriendo los usuarios durante la realización de la prueba ya que el test de Tukey determina que los Órdenes 1 (ABO - PRO - PRE) y 2 (ABO - PRE - PRO) tienen un comportamiento diferente al Orden 5 (PRE - ABO - PRO).

Realmente, el experimento realizado no nos permite lograr explicar más a fondo esta situación y por ello se vuelve necesario continuar la investigación para poder explicar estos resultados.

En el caso de la variable número de clicks la situación vuelve a ser compleja. Como se comenta en la sección 3.12.1.1, aquí los órdenes que son diferentes son

el 2 y el 5. En este caso ni siquiera interviene el orden 1. Nuevamente se vuelve necesario continuar la investigación para conseguir analizar con mayor eficacia qué está sucediendo en este punto.

CAPÍTULO 5. CONCLUSIONES Y TRABAJOS FUTUROS

5.1. Conclusiones

Al inicio de la investigación, sección 1.3, se marcaron cinco objetivos.

Referente al objetivo *Estudiar en profundidad la literatura y los trabajos existentes sobre cómo realizar evaluaciones empíricas para medir atributos de usabilidad y determinar el estado del arte de las investigaciones* no se ha encontrado en la literatura ningún experimento con las características del llevado a cabo, por lo que la investigación es original y relevante.

El objetivo *Formalizar el análisis del estudio para obtener las evidencias necesarias para construir la investigación y diseñar el experimento* ha sido superado en cuanto se ha llevado a cabo el experimento, los experimentos si se añade el experimento piloto, desde el análisis del problema hasta el diseño y ejecución del experimento.

Los tres últimos objetivos *Estudiar empíricamente la influencia de entre la presencia o la ausencia del mecanismo Preferencias | Abortar Operación | Retroalimentación de Progreso en la usabilidad de la aplicación* se cubren durante el experimento.

Realmente, lo importante es lo que se desprende del estudio empírico para estos tres mecanismos de usabilidad. Los resultados son concluyentes para la variable respuesta satisfacción pues la presencia de los tres mecanismos mencionados hace que la satisfacción aumente. La eficiencia nunca empeora cuando la aplicación presenta alguno de los mecanismos pero, tal y como se expone en el cuarto capítulo, en alguno de los casos se ha encontrado que el experimento utilizado hacía que el factor orden fuese determinante.

En el caso del mecanismo *Abortar Operación* sí se observa una mejora significativa en la eficiencia del usuario (tanto en su rapidez como en el nivel de interacción con la aplicación) en la realización de la tarea cuando este mecanismo está presente. En el caso del mecanismo *Retroalimentación de Progreso* la eficiencia sólo es significativa cuando se mide como el nivel de interacción del usuario con la aplicación. Sin embargo, para el mecanismo *Preferencias* se observa que el orden de realización de las tareas es un factor decisivo.

Es por ello, que de este estudio se desprende que la incorporación de los mecanismos de usabilidad estudiados *Abortar Operación*, *Retroalimentación de Progreso* y *Preferencias* mejora la aplicación web que se esté creando, ya que los usuarios están más satisfechos y realizan las tareas con eficiencia.

5.2. Limitaciones del Trabajo

Aunque el experimento ha sido realizado por 100 sujetos, 20 más de los exigidos por cuestiones matemáticas, éstos tienen características similares: la mayoría son estudiantes de Bachillerato, con conocimientos y comportamientos similares. Esto puede ser una limitación, ya que los resultados pueden no ser representativos para una muestra real aunque, hoy en día, tal y como se demuestra en el cuestionario de familiaridad, la gente joven cada vez está más acostumbrada al uso de Internet y de aplicaciones que permitan las compras online. Ver sección 3.9 donde se comenta la validez, tanto interna como externa, del experimento realizado.

También existen limitaciones relacionadas con el tipo de aplicación. El experimento se hace en una aplicación web con alto contenido interactivo. Este tipo de aplicación tiene una caracterización y comportamiento diferente en relación con otros tipos de aplicaciones, lo que puede suponer que los resultados obtenidos en este caso no sean fácilmente extrapolables a otras situaciones.

5.3. Trabajos Futuros

Con base a los resultados y aportes de este estudio se abren nuevas líneas de investigación que permitan afianzar los conocimientos adquiridos fruto de este trabajo. Los trabajos futuros que se considerarán son los siguientes:

- Estudiar nuevas variables respuesta, tales como la eficacia o la eficiencia relativa.
- Realizar réplicas del experimento con sujetos que no provengan de una muestra tan específica como la utilizada en esta investigación.
- Modificar ligeramente el experimento de forma que el orden deje de ser un factor relevante en el mismo.
- Realizar el experimento en aplicaciones de tienda online con otros productos para generalizar los resultados en aplicaciones *e-commerce*.
- Realizar el experimento con una aplicación que no sea una tienda para estudiar el comportamiento y la variación de los resultados entre los dos tipos de aplicaciones.

REFERENCIAS

Acuña, S.T., Castro, J.W. A Systematic Mapping Study on the Open Source Software Development Process, in *Proceedings of the International Conference on Evaluation and Assessment in Software Engineering (EASE'12)*, pp.1-5, 2012.

Acuña, S.T., Gómez, M., Juristo, N. Towards Understanding the Relationship Between Team Climate and Software Quality - A Quasi-Experimental Study. *Empirical Software Engineering*, Springer, vol. 13, no. 4, pp. 401-434, 2008.

Andreasen, M.S., Nielsen, H.V., Schröder, S.O., Stage, J. What happened to Remote Usability Testing? An Empirical Study of Three Methods, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*, pp. 1405-1411, ACM, 2007.

Bass, L., John, B.E. Linking Usability to Software Architecture Patterns through General Scenarios. *Journal of Systems and Software*, vol. 66, no. 3, pp. 187-197, 2003.

Bass, L., John, B.E., Kates, J. *Achieving Usability through Software Architecture*. Technical Report CMU/SEI-TR-2001-005, Software Engineering Institute, Carnegie Mellon University, Pittsburg, 2001.

Batthey, J. IBM's Redesign Results in a Kinder, Simpler Web Site, 2001. Available at: http://interface.free.fr/Archives/IBM_redesign_results.pdf

Black, J. Usability Is Next to Profitability, 2002. Available at: http://www.businessweek.com/technology/content/dec2002/tc2002124_2181.htm

Bolchini, D., Does Branding Need Web Usability? A Value-Oriented Empirical Study, in *Proceedings of the 12th IFIP TC13 Conference on Human-Computer Interaction*, Springer, pp. 652-665, 2009.

Campbell, D.T., Stanley, J.C. *Experimental and Quasi-Experimental Designs for Research*. Boston, MA: Houghton Mifflin Company, 1963.

Cassino, R., Tucci, M., Vitiello, G. Empirical Validation of an Automatic Usability Evaluation Method. *Information and Software Technology*, Elsevier, vol. 28, pp. 1-22, 2015.

Castro, V., Genero, M., Marcos, E., Piattini, M. Empirical Study to Assess Whether the Use of Routes Facilitates the Navigability of Web Information Systems. *IET Software*, vol. 5, no. 6, pp. 528-542, 2011.

Cochran, W.C., Cox, G.M. *Diseños Experimentales*. Distrito Federal, México: Trillas, 1980.

Donahue, G.M. Usability and the Bottom Line. *IEEE Software*, vol. 18, no. 1, pp. 31-37, 2001.

Fang, X., Holsapple, C.W. An Empirical Study of Web Site Navigation Structures' Impacts on Web Site Usability. *Decision Support Systems*, vol. 43, no. 2, pp. 476-491, 2007.

Ferré, X., Juristo, N., Windl, H., Constantine, L. Usability Basics for Software Developers. *IEEE Software*, vol. 18, no. 1, pp. 22-29, 2001.

- Field, A., Miles, J., Field, Z. *Discovering Statistics Using R*. London, United Kingdom: SAGE, 2012.
- Frøkjær, E., Hertzum, M., Horbæk, K. Measuring Usability: Are Effectiveness, Efficiency and Satisfaction Really Correlated?, in *Proceedings of Human Factor in Computer Systems (CHI'00)*, pp. 345-352, ACM, 2000.
- Horbæk, K., Frøkjær, E. Two Psychology-Based Usability Inspection Techniques Studied in a Diary Experiment, in *Proceedings of the Third Nordic Conference on Human-Computer Interaction (NordiCHI'04)*, pp. 3-12, ACM, 2004.
- ISO/IEC 25010. Systems and Software Engineering. System and Software Product. Quality Requirements and Evaluation (SQuaRE). System and Software Quality Models, 2010.
- ISO 9241-11. Ergonomics Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability, 1998.
- Juristo, N., Moreno, A.M., Sánchez-Segura, M.I. Analysing the Impact of Usability on Software Design. *Journal of System and Software*, vol. 80, no. 9, pp. 1506-1516, 2007a.
- Juristo, N., Moreno, A.M., Sánchez-Segura, M.I. Guidelines for Eliciting Usability Functionalities Software Engineering. *IEEE Transactions on Software Engineering*, vol. 33, no. 11, pp. 744-758, 2007b.
- Juristo, N., Moreno, A.M. Basics of Software Engineering Experimentation. Boston, MA: Springer, 2001.
- Kitchenham, B.A. *Guidelines for Performing Systematic Literature Reviews in Software Engineering Version 2.3*. Technical Report S.o.C.S.a.M. Software Engineering Group, Keele University and Department of Computer Science University of Durham, 2007.
- Kitchenham, B.A. *Procedures for Performing Systematic Reviews*. Technical Report TR/SE-0401 and NICTA Technical Report 0400011T.1, Tech. Rep., Keele University, 2004.
- Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M. Systematic Mapping Studies in Software Engineering, in *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE'08)*, pp. 71-80, 2007.
- Piccioni, M., Furia, C.A., Meyer, B. An Empirical Study of API Usability, in *Proceedings of the ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, pp. 56-62, 2013.
- Rodríguez, F.D. Obtención y Uso de Patrones para la Implementación de Funcionalidades de Usabilidad en Aplicaciones Web. Tesis Doctoral. Directoras: Acuña, S.T., Juristo, N. Universidad Politécnica de Madrid. 2014. Available at: <http://oa.upm.es/33528/>
- Rodríguez, F.D. Programming Patterns for Usability Functionalities. *GrISE - Grupo de Investigación en Ingeniería del Software Empírica*, Universidad Politécnica de Madrid. 2010. Available at: <http://www.grise.upm.es/sites/extras/7/>
- Sauro, J., Kindlund, E. A Method to Standardize Usability Metrics Into a Single Score, in *Proceedings of the SIGCHI Conference on Human Computing System*, pp. 401-409, ACM, 2005.

Tullis, T.S., Stetson, I.N. A Comparison of Questionnaires for Assessing Website Usability, in *Proceedings of the Usability Professionals Association (UPA) Conference*, pp. 7-11, 2004.

Winter, J., Rönkkö, K. SPI Success Factors within Product Usability Evaluation. *Journal of Systems and Software*, vol. 83, pp. 2059-2072, 2010.

GLOSARIO

ANOVA: (ANOVA, ANalysis Of VAriance según la terminología inglesa) el análisis de la varianza, en estadística, es una técnica estadística de contraste de hipótesis. Su utilidad radica en que permite comparar las medias de más de dos muestras, situación en la que no es correcto utilizar un análisis basado en el test *t* de *Student*.

Asimetría: en estadística, las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin necesidad de realizar su representación gráfica. En este estudio se utiliza el *Coefficiente de Asimetría de Fisher* que utiliza el tercer momento estándar y es la medida de asimetría más utilizada.

Curtosis: en teoría de la probabilidad y estadística, la curtosis es una medida de la forma. Las medidas de curtosis tratan de estudiar la proporción de la varianza que se explica por la combinación de datos extremos respecto a la media en contraposición con datos poco alejados de la misma.

Cuasi experimento: diseños de investigación experimentales en los cuales los sujetos o grupos de sujetos de estudio no están asignados aleatoriamente a los tratamientos.

Diseño factorial completo: diseño que consta de dos o más factores, cada uno de los cuales presenta distintos niveles. Las unidades experimentales cubren todas las posibles combinaciones de todos los niveles de todos los factores.

Diseño factorial fraccionado: diseño que consta de dos o más factores, cada uno de los cuales presenta distintos niveles. Las unidades experimentales están asignadas a una fracción de las posibles combinaciones de los niveles de los factores.

Estudio empírico: modelo de investigación científica que se basa en la experimentación y la lógica empírica que junto con la observación de fenómenos y su análisis estadístico, es el más usado en el campo de las ciencias sociales y las ciencias naturales.

Evidencia empírica: información que es adquirida a través de la observación o de la experimentación.

Experimento verdadero: diseños de investigación experimentales en los cuales los sujetos o grupos de sujetos de estudio están asignados aleatoriamente a los tratamientos.

ANEXOS

Anexo A. Evolución del Diseño

Como se comentaba en la sección 3.2 del cuerpo de este documento, el diseño del experimento verdadero aplicado en Galicia es fruto de una serie de propuestas de diseño que han ido cambiando a lo largo de toda la investigación. Este anexo pretende recoger los cambios que ha ido sufriendo el diseño del experimento.

A.1. Diseño Full Factorial

Inicialmente, el diseño experimental escogido fue un diseño Full Factorial, tomando como factor el mecanismo de usabilidad (MU) con sus dos niveles (presente=1, o ausente=0) para cada uno de los mecanismos Abortar Operación (ABO), Retroalimentación de Progreso (PRO) y Preferencias (PRE). Por tanto, en este caso se tienen ocho agrupaciones que resultan de elevar los dos niveles a los tres tratamientos.

Como cada sujeto debe realizar varias tareas, es necesario analizar las secuencias de ejecución de esas tareas asociadas a cada uno de los mecanismos de usabilidad. De aquí, una idea inicial sustenta la construcción de varios escenarios con tareas con el fin de considerar sólo las secuencias que resulten naturales y descartar las restantes. Hay que recordar que un escenario está compuesto por varias tareas que un usuario debe realizar pero, sin embargo, se debe cuidar que cada uno de los escenarios utilicen una secuencia de ejecución diferente. La Tabla 41 recoge la matriz con los grupos, donde un cero indica que no está presente el mecanismo y un uno que está presente el mecanismo en cuestión.

La secuencia de ejecución de tareas en un escenario conforma lo que se denomina orden de exposición del mecanismo de usabilidad y están numeradas del 1 al 6 tal y como se puede observar en la Tabla 42. Un escenario puede ejecutar el segundo orden indicando que primeramente se pedirá al participante que realice la tarea ABO, seguidamente la tarea PRE y, por último, la tarea PRO. Otro escenario podría ser el cuarto y así sucesivamente.

La matriz de asignación de agrupaciones correspondiente a la combinación entre la matriz de grupos y el orden de exposición del mecanismo de usabilidad se muestra en la Tabla 43. Se observa como, por ejemplo, tomando el orden de exposición O1 para el grupo B, el escenario debe solicitar la ejecución de tareas donde ABO esté como primer mecanismo inactivo, seguido de PRO inactivo y finalmente PRE activo.

| | ABO | PRO | PRE |
|---|-----|-----|-----|
| A | 0 | 0 | 0 |
| B | 0 | 0 | 1 |
| C | 0 | 1 | 0 |
| D | 0 | 1 | 1 |
| E | 1 | 0 | 0 |
| F | 1 | 0 | 1 |
| G | 1 | 1 | 0 |
| H | 1 | 1 | 1 |

Tabla 41: Matriz de grupos Full Factorial

| | Primero | Segundo | Tercero |
|----|---------|---------|---------|
| O1 | ABO | PRO | PRE |
| O2 | ABO | PRE | PRO |
| O3 | PRO | PRE | ABO |
| O4 | PRO | ABO | PRE |
| O5 | PRE | ABO | PRO |
| O6 | PRE | PRO | ABO |

Tabla 42: Matriz con el orden de exposición de cada MU del diseño Full Factorial

| | | Primero | | | Segundo | | | Tercero | | |
|---|----|---------|-----|-----|---------|-----|-----|---------|-----|-----|
| | | ABO | PRO | PRE | ABO | PRO | PRE | ABO | PRO | PRE |
| A | O1 | 0 | | | | 0 | | | | 0 |
| | O2 | 0 | | | | | 0 | | 0 | |
| | O3 | | 0 | | | | 0 | 0 | | |
| | O4 | | 0 | | 0 | | | | | 0 |
| | O5 | | | 0 | 0 | | | | 0 | |
| | O6 | | | 0 | | 0 | | 0 | | |
| B | O1 | 0 | | | | 0 | | | | 1 |
| | O2 | 0 | | | | | 1 | | 0 | |
| | O3 | | 0 | | | | 1 | 0 | | |
| | O4 | | 0 | | 0 | | | | | 1 |
| | O5 | | | 1 | 0 | | | | 0 | |
| | O6 | | | 1 | | 0 | | 0 | | |
| H | O1 | 1 | | | | 1 | | | | 1 |
| | O2 | 1 | | | | | 1 | | 1 | |
| | O3 | | 1 | | | | 1 | 1 | | |
| | O4 | | 1 | | 1 | | | | | 1 |
| | O5 | | | 1 | 1 | | | | 1 | |
| | O6 | | | 1 | | 1 | | 1 | | |

Tabla 43: Matriz de asignación de agrupaciones con órdenes de exposición del diseño Full Factorial

A.2. Primera Versión del Diseño Ortogonal

Como segunda aproximación, y debido a que el diseño Full Factorial presentado requiere tener muchas unidades experimentales, especialmente sujetos, tras varias reuniones del equipo de investigación se decidió intentar un diseño ortogonal en el cual se fraccionase la matriz factorial de grupos en un subconjunto que requiriese menos sujetos pero que no excluyese opciones altamente significativas. En la Tabla 44 se visualiza la matriz ortogonal de grupos. También se analizó la posibilidad de utilizar una matriz de cuadrados latinos para el orden de exposición de los distintos mecanismos de usabilidad donde cada grupo aparece exactamente una vez en cada fila y en cada columna. Esta matriz se describe en la Tabla 45. Para completar este diseño, se presenta la matriz de asignación de grupos resultante de combinar la matriz ortogonal de grupos y el orden por cuadrados latinos. Dicha matriz está contenida en la Tabla 46. Con estos cambios se consigue reducir significativamente el coste de diseño factorial manteniendo la ortogonalidad y evitando la correlación entre los factores.

| | ABO | PRO | PRE |
|---|-----|-----|-----|
| A | 0 | 0 | 1 |
| B | 0 | 1 | 0 |
| C | 1 | 0 | 0 |
| D | 1 | 1 | 1 |

Tabla 44: Matriz de grupos versión 1 del Diseño Ortogonal

| | Primero | Segundo | Tercero |
|----|---------|---------|---------|
| O1 | ABO | PRO | PRE |
| O2 | PRE | ABO | PRO |
| O3 | PRO | PRE | ABO |

Tabla 45: Matriz con el orden de exposición por Cuadrados Latinos

| | | Primero | | | Segundo | | | Tercero | | |
|---|----|---------|-----|-----|---------|-----|-----|---------|-----|-----|
| | | ABO | PRO | PRE | ABO | PRO | PRE | ABO | PRO | PRE |
| A | O1 | 0 | | | | 0 | | | | 1 |
| | O2 | | | 1 | 0 | | | | 0 | |
| | O3 | | 0 | | | | 1 | 0 | | |
| B | O1 | 0 | | | | 1 | | | | 0 |
| | O2 | | | 0 | 0 | | | | 1 | |
| | O3 | | 1 | | | | 0 | 0 | | |
| C | O1 | 1 | | | | 0 | | | | 0 |
| | O2 | | | 0 | 1 | | | | 0 | |
| | O3 | | 0 | | | | 0 | 1 | | |
| D | O1 | 1 | | | | 1 | | | | 1 |
| | O2 | | | 1 | 1 | | | | 1 | |
| | O3 | | 1 | | | | 1 | 1 | | |

Tabla 46: Matriz de asignación de agrupaciones con órdenes de exposición

A.3. Segunda Versión del Diseño Ortogonal

Siguiendo con el diseño ortogonal, se realizó un análisis de cada uno de los mecanismos de usabilidad que se pretendían estudiar. En este análisis se tuvo en cuenta los comportamientos de los mecanismos definidos en los artefactos de diseño y programación de [Rodríguez, 2010]. En este sentido, un comportamiento reflejado del mecanismo de usabilidad Preferencias hace hincapié en la personalización de la interfaz haciendo que los efectos de dichos mecanismo se propague a lo largo de toda la aplicación y se implementa de manera independiente a cualquier otra funcionalidad. En estas condiciones es razonable pensar que pedir la tarea asociada a PRE al final de la secuencia carece de sentido. Es por ello que el O1 de la Tabla 45 no puede ser ejecutado, es decir, la matriz de cuadrados latinos deberá ser reemplazado por otra en la cual la secuencia de tareas no termine en PRE. Analizando el orden de exposición definido en el diseño factorial se descartaron varios órdenes de ejecución, resultando únicamente válidos {O2, O3, O5, O6}. La Tabla 47 muestra los órdenes válidos y los descartados.

| | Primero | Segundo | Tercero | Validez |
|-----------|---------|---------|---------|-----------|
| O1 | ABO | PRO | PRE | No válida |
| O2 | ABO | PRE | PRO | Correcta |
| O3 | PRO | PRE | ABO | Correcta |
| O4 | PRO | ABO | PRE | No válida |
| O5 | PRE | ABO | PRO | Correcta |
| O6 | PRE | PRO | ABO | Correcta |

Tabla 47: Órdenes de exposición válidos de la segunda versión del Diseño Ortogonal

La matriz de asignación de grupos resultante en este diseño se puede visualizar en la Tabla 48.

| | | Primero | | | Segundo | | | Tercero | | |
|----------|-----------|---------|-----|-----|---------|-----|-----|---------|-----|-----|
| | | ABO | PRO | PRE | ABO | PRO | PRE | ABO | PRO | PRE |
| A | O2 | 0 | | | | | 1 | | 0 | |
| | O3 | | 0 | | | | 1 | 0 | | |
| | O5 | | | 1 | 0 | | | | 0 | |
| | O6 | | | 1 | | 0 | | 0 | | |
| B | O2 | 0 | | | | | 0 | | 1 | |
| | O3 | | 1 | | | | 0 | 0 | | |
| | O5 | | | 0 | 0 | | | | 1 | |
| | O6 | | | 0 | | 1 | | 0 | | |
| C | O2 | 1 | | | | | 0 | | 0 | |
| | O3 | | 0 | | | | 0 | 1 | | |
| | O5 | | | 0 | 1 | | | | 0 | |
| | O6 | | | 0 | | 0 | | 1 | | |
| D | O2 | 1 | | | | | 1 | | 1 | |
| | O3 | | 1 | | | | 1 | 1 | | |
| | O5 | | | 1 | 1 | | | | 1 | |
| | O6 | | | 1 | | 1 | | 1 | | |

Tabla 48: Matriz de asignación de agrupaciones con órdenes de exposición

A.4. Diseño Between Subjects del Experimento Piloto

Avanzando en el diseño y tras realizar un análisis más profundo del comportamiento del mecanismo de usabilidad Preferencias se encontró que, tomando las secuencias válidas en la Tabla 47, existen ciertos órdenes que, aunque sean válidos, no parecen ser secuencias naturales para un usuario. Concretamente, si tomamos el grupo A con orden O2 de la Tabla 48 solicitaremos al sujeto la tarea ABO, después intentará personalizar su interfaz, PRE, y por último, realizará la búsqueda de un ítem o artículo, PRO. En estas condiciones, cuando el usuario inicia la tarea, lo hará con una interfaz no agradable para la tarea ABO. Después, con la tarea PRE, el usuario obtendría una interfaz potencialmente agradable para realizar posteriormente la tarea PRO bajo estas últimas condiciones. Entonces, se notan ciertas situaciones confusas, ya que el usuario podría desorientarse al conmutar de interfaz durante la realización de las tareas. Además, los cuestionarios de satisfacción se verían sesgados por la probable confusión que el usuario entienda entre la interfaz personalizada y la interfaz no agradable.

Por otra parte, como el mecanismo PRE afecta a toda la aplicación, el comportamiento de la nueva interfaz se propaga para la tarea siguiente, es decir, en este caso concreto la tarea PRO se realizaría con una interfaz completamente distinta que la tarea ABO.

Atendiendo a que el MU-PRE afecta a toda la aplicación, es razonable pensar que sólo las secuencias O5 y O6 de la Tabla 47 son lógicas para la realización de las tareas. De esta forma se decide eliminar la probable confusión que el usuario podría tener cuando alterna entre interfaces distintas. Otra consideración del mecanismo PRE es que la tarea asociada a este mecanismo no es en esencia una tarea porque se le pide al usuario que use las preferencias. En este sentido, se propone una tarea ficticia asociada a la tarea propia de PRE para medir el impacto de las preferencias (mejora en la usabilidad) en la realización de una tarea distinta a las tareas ABO y PRO.

Bajo estas condiciones se modifica el diseño alterando la matriz ortogonal de grupos, donde sustituimos la última fila – grupo D – por valores ceros (Tabla 49), los cuales representarán el conjunto de sujetos del grupo de control, GC, y los grupos A, B y C serán los grupos de testeo, GT.

La idea con este nuevo cambio es que el análisis se basaba en la evaluación del escenario, cuyo mecanismo de usabilidad esté presente contra el escenario donde el mecanismo esté ausente. Esta nueva matriz de grupos cumple el diseño *between subjects* (BS) donde las evaluaciones de los distintos mecanismos se realizan por filas (A contra D para PRE, B contra D para PRO y C contra D para ABO). Además, con este diseño se ha reducido el orden de exposición sólo a las secuencias que se consideran lógicas para la realización de las tareas (Tabla 50). La idea de las secuencias lógicas es que sea un bloque continuo de ejecución de tareas, es decir, tareas lógicamente ordenadas.

La matriz de asignación de grupos resultante de combinar la matriz BS de grupos y los órdenes lógicos se proporciona en la Tabla 51.

| | ABO | PRO | PRE |
|---|-----|-----|-----|
| A | 0 | 0 | 1 |
| B | 0 | 1 | 0 |
| C | 1 | 0 | 0 |
| D | 0 | 0 | 0 |

Tabla 49: Matriz de grupos *Between Subject*

| | Primero | Segundo | Tercero |
|----|---------|---------|---------|
| O1 | PRE | ABO | PRO |
| O2 | PRE | PRO | ABO |

Tabla 50: Matriz con el orden de exposición con secuencias lógicas

| | | Primero | | | Segundo | | | Tercero | | |
|---|----|---------|-----|-----|---------|-----|-----|---------|-----|-----|
| | | ABO | PRO | PRE | ABO | PRO | PRE | ABO | PRO | PRE |
| A | O1 | | | 1 | 0 | | | | 0 | |
| | O2 | | | 1 | | 0 | | 0 | | |
| B | O1 | | | 0 | 0 | | | | 1 | |
| | O2 | | | 0 | | 1 | | 0 | | |
| C | O1 | | | 0 | 1 | | | | 0 | |
| | O2 | | | 0 | | 0 | | 1 | | |
| D | O1 | | | 0 | 0 | | | | 0 | |
| | O2 | | | 0 | | 0 | | 0 | | |

Tabla 51: Matriz de asignación de agrupaciones con órdenes de exposición

El resumen del diseño es el siguiente:

- Diseño: *between subjects*.
- 4 Grupos: PRE (A), PRO (B), ABO (C), None (D).
- 1 Factor: mecanismo de usabilidad presente o ausente.
- Variable de bloque: secuencia de realización de las tareas (no se analizará pero deberá ser tenido en cuenta de tal modo que se garantice que el orden de ejecución de las tareas no influya en el resultado).
- Asignación de agrupaciones: aleatorio puro. Cada sujeto es asignado aleatoriamente en una de las agrupaciones, asegurando que se cumpla el balanceo de la cantidad de miembros por grupo, Tabla 51. Se esperan aproximadamente 10 sujetos por grupo para la realización del test ANOVA.
- Tareas: se tiene una tarea por mecanismo. Además, existirá una tarea ficticia para evaluar el impacto del MU-PRE.
 - Tarea PRE: Explorar y jugar con la interfaz para cambiarla, si se quiere.
 - Tarea PRE Ficticia (PRE-F): Buscar información del plazo de devolución de cualquier artículo o ítem.
 - Tarea ABO: Cancelar las modificaciones de un carrito.
 - Tarea PRO: Buscar un artículo en la tienda y agregarlo al carrito.

El orden de las tareas está determinado por el orden de exposición reflejado en la Tabla 50.

- Cuestionarios post tarea: como el tiempo de realización de las tareas es breve es preferible hacer un cuestionario final de todos los mecanismos, en lugar de interrumpir las tareas con un cuestionario relativamente corto. Se incluye un único cuestionario final con preguntas sobre todos los mecanismos.

Éste es el diseño utilizado en el experimento piloto detallado en el Anexo B.

Anexo B. Experimento Piloto

B.1. Introducción

Antes de realizar el experimento verdadero final en Galicia se tuvo la oportunidad de hacer una prueba piloto que tenía como objetivo recoger información para evaluar si el diseño que se estaba utilizando era correcto o necesitaba cambios.

Con este fin, el 23 de noviembre, los estudiantes de la asignatura Desarrollo de Sistemas Interactivos del Máster Universitario en Ingeniería Informática de la Universidad Autónoma de Madrid, participaron en el experimento en una sesión que duró aproximadamente 30 minutos.

Los objetivos y las hipótesis del experimento vienen fijados por los objetivos e hipótesis de investigación del trabajo y han sido mencionadas en las secciones 3.3 y 3.4, respectivamente. Las variables respuesta estudiadas también son generales a toda la investigación y se detallan en la sección 3.5.

En cuanto al diseño experimental se utilizó el diseño *between subjects* descrito en la sección A.4. Respecto al método de instrumentación se proponen, como en el experimento final, cuatro tareas (las propias de cada uno de los mecanismos de usabilidad y la tarea ficticia para el mecanismo Preferencias), se puede consultar la sección 3.7 y el Anexo C para obtener una visión más completa de las mismas. Por último, mencionar que el procedimiento de recolección de datos es idéntico al detallado en la sección 3.8 para el experimento verdadero final.

B.2. Análisis de Datos

El método estadístico utilizado para la realización del análisis del experimento piloto se divide en los siguientes pasos:

1. *Datos Brutos*: datos contenidos en las tablas de métricas bajo las indicaciones detalladas en la sección 3.8 de este documento en el caso de la eficiencia. Datos contenidos en las tablas de los cuestionarios de las tareas en el caso de la satisfacción.
2. *Análisis descriptivo*: como primer paso para entender los datos se efectuará un análisis básico descriptivo. La importancia de este análisis radica en que informa de cómo es la muestra de la que se dispone. Entre otros, se calcularán los siguientes estadísticos descriptivos: mínimo, máximo, media, mediana, desviación típica, coeficiente de asimetría, rango intercuartílico, etc. Para completar este análisis descriptivo se realizarán gráficos de caja que ayuden a comprender la forma de los datos.
3. *Extracción de valores imposibles*: tras analizar los estadísticos descriptivos y para mantener la coherencia de los resultados se han decidido eliminar los valores imposibles. Se consideran imposibles aquellos valores que por las

características del experimento determinan que la tarea o bien no se ha hecho o bien se ha hecho de forma incorrecta, por ejemplo, eliminar aquellos sujetos que han realizado una tarea con 0 clicks o en 0 segundos.

En cuanto a los datos de satisfacción, no tiene sentido eliminar valores imposibles, pues los valores siempre están controlados por la mínima y máxima puntuación posible en los cuestionarios.

4. *Prueba de normalidad de residuos*: para continuar el análisis se realiza la prueba de normalidad. Concretamente se realiza una de las pruebas más utilizadas para este fin, la prueba de *Shapiro-Wilk*.
5. *Prueba de homogeneidad de varianzas*: además, previamente a la realización del test ANOVA es necesario asegurar la homogeneidad de las varianzas. Para ello se utiliza el test de *Levene*, un test estadístico utilizado para evaluar la igualdad de las varianzas de una variable calculada para dos o más grupos.
6. *Test ANOVA*: por último, se realizará un test ANOVA para comprobar si se puede aceptar o rechazar la hipótesis nula en cada caso. Concretamente se realizará un test ANOVA Factorial Fraccionado. La Tabla 52 muestra los distintos elementos considerados en el test ANOVA independientemente del mecanismo de usabilidad que se pretenda analizar.

| Factor | Grado | Valores posibles |
|--------|-----------|-----------------------|
| PA | Principal | Presencia - Ausencia |
| Grupo | Segundo | A – B – C – D |
| Orden | Tercero | 1 – 2 – 3 – 4 – 5 – 6 |

Tabla 52: Elementos considerados en los test ANOVA para el experimento piloto

Importante: por las características que presenta este experimento los datos deben ser analizados de forma agrupada, es decir, cada una de las variables consideradas (número de clicks, tiempo y valor medio de las respuestas en los cuestionarios) se suman para las tres tareas estudiadas (ABO, PRO y PRE-F). El test ANOVA realizado es el mismo que en el caso del experimento verdadero, es decir, un ANOVA Factorial Fraccionado de acuerdo a la Tabla 52.

B.3. Análisis Descriptivo

B.3.1. Eficiencia

En la Tabla 53 se presentan los resultados referidos al atributo Eficiencia. En cuanto a la variable número de clicks se observa como la media es menor cuando los mecanismos están presentes y mayor cuando están ausentes. El comportamiento es similar para todos los cuartiles salvo en el primer cuartil donde en ambas muestras coinciden. De la misma forma, la desviación típica es mayor en el caso de la ausencia de los mecanismos. Todos los coeficientes de asimetría y Curtosis son positivos.

En el caso de la variable tiempo, también se observa que la media, el primer cuartil, el tercer cuartil y el cuarto cuartil son menores cuando los mecanismos están activos. La desviación típica toma valores muy similares aunque un poco superior en el caso de la presencia de los mecanismos. En este caso también son positivos los coeficientes de asimetría y Curtosis aunque más próximos a cero.

| AGRUPADO | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|----------|--------|--------|--------|---------|------|-------|-------|-------|--------|--------|----------|----------|----------|
| Clicks P | 7 | 14 | 9.167 | 8.5 | 3 | 7 | 8.5 | 10 | 14 | 7.28 | 7.76 | 1.15 | 0.86 |
| Clicks A | 1 | 33 | 12.26 | 10 | 7.5 | 7 | 10 | 14.5 | 33 | 8.48 | 71.92 | 1.19 | 1.25 |
| Tiempo P | 30.396 | 107.31 | 63.096 | 67.39 | 28.2 | 42.54 | 67.39 | 70.72 | 107.31 | 28.035 | 785.96 | 0.401 | 0.10 |
| Tiempo A | 32.688 | 138.36 | 73.049 | 66.67 | 31.3 | 55.27 | 66.67 | 86.53 | 138.36 | 27.540 | 758.47 | 0.937 | 0.93 |

Tabla 53: Estadísticos descriptivos para la Eficiencia en el experimento piloto

En la Figura 13 se presenta el diagrama de cajas para la variable número de clicks. Se observa como las muestras son diferentes, aunque tanto el primer cuartil como la mediana toman valores muy parecidos. Destaca que la muestra que dispone de los mecanismos activos no tiene bigote inferior mientras que la muestra de los sujetos que no disponen de los mecanismos sí lo tiene, y además, bastante pronunciado. En este diagrama se observa un único valor atípico en la muestra Ausencia.

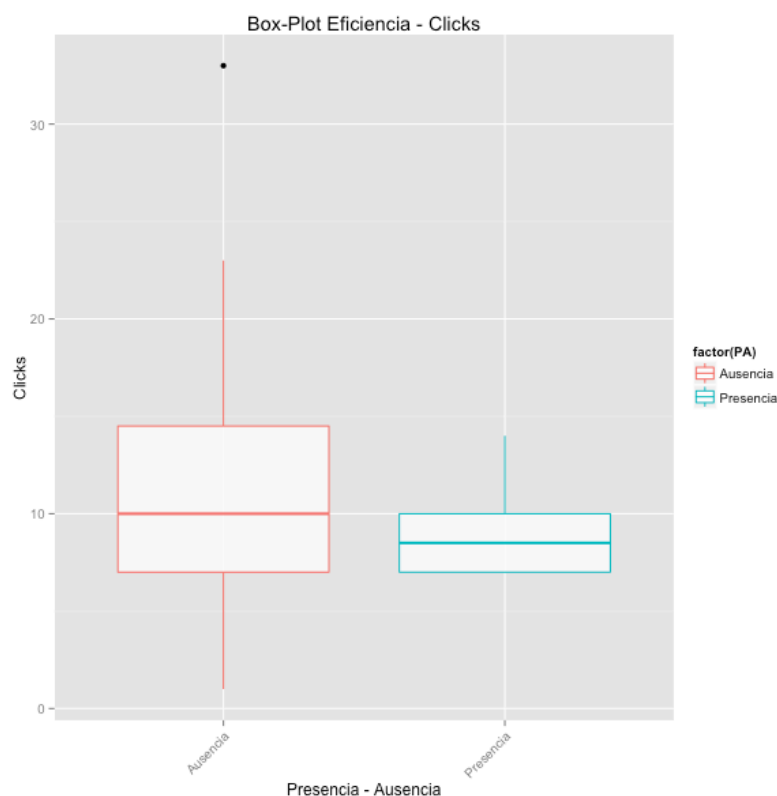


Figura 13: Diagrama de cajas, datos agrupados, para la variable N° de Clicks en el experimento piloto

En la Figura 14 se presenta el diagrama de cajas para la variable tiempo. Se observa como las muestras no son simétricas. Si bien toman un valor similar en la mediana, la

muestra de los sujetos que no disponían de los mecanismos de usabilidad presenta una caja donde dicha mediana está más centrada frente a la muestra de los sujetos que sí disponían del mecanismo, donde que la mediana está claramente en la parte alta de la caja, casi coincidiendo con el tercer cuartil. Ambas muestras presentan bigotes superiores e inferiores y solamente hay un valor atípico en la muestra de los sujetos con los mecanismos inactivos.

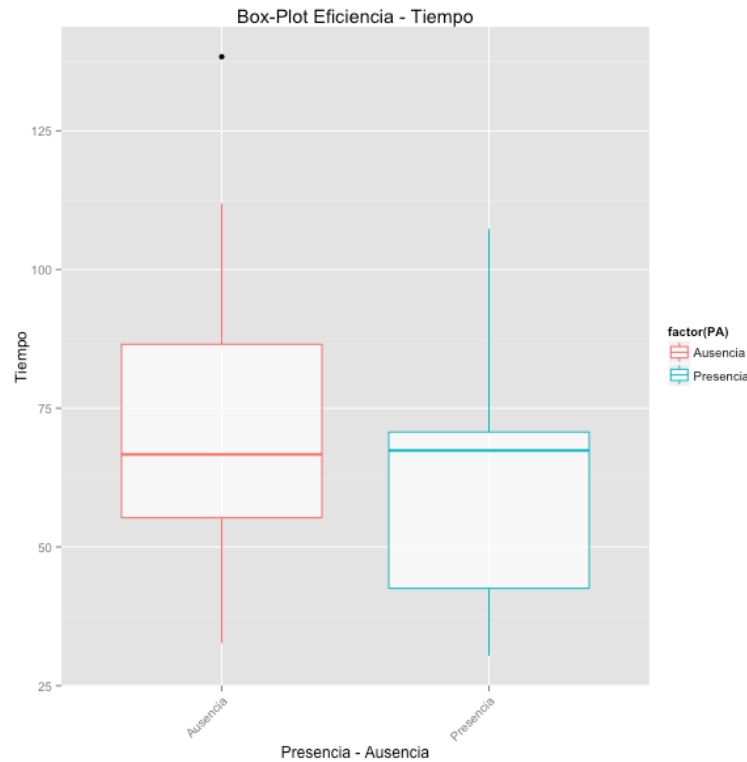


Figura 14: Diagrama de cajas, datos agrupados, para la variable Tiempo en el experimento piloto

B.3.2. Satisfacción

En la Tabla 54 aparecen los estadísticos descriptivos para la variable Satisfacción. Se observa como los resultados son bastante similares en muchos de ellos, por ejemplo, casi coinciden las medias, medianas, cuartiles, desviación típica, mínimo, máximo y varianza. Los coeficientes de asimetría también tomas valores similares ambos positivos mientras que los coeficientes de Curtosis son negativos.

| AGRUPADO | Mín. | Máx. | Media | Mediana | IQR | Q1 | Q2 | Q3 | Q4 | SD | Varianza | Skewness | Kurtosis |
|----------|------|------|-------|---------|------|------|------|------|------|------|----------|----------|----------|
| Valor P | 1 | 4 | 2.55 | 2.5 | 1.25 | 2 | 2.5 | 3.25 | 4 | 0.83 | 0.70 | 0.096 | -0.85 |
| Valor A | 1 | 4.33 | 2.40 | 2.33 | 1 | 1.83 | 2.33 | 2.83 | 4.33 | 0.86 | 0.73 | 0.16 | -0.40 |

Tabla 54: Estadísticos descriptivos para la Satisfacción en el experimento piloto

En la Figura 15 está representado el diagrama de cajas de la variable Satisfacción. Para ambas muestras la forma de la caja es similar aunque, la barra de la mediana aparece más centrada en el caso de la ausencia del mecanismo. Como era de esperar ambas muestras tienen bigotes que terminan en los mismos puntos. Nuevamente sólo se presenta un valor atípico en el caso de la ausencia del mecanismo.

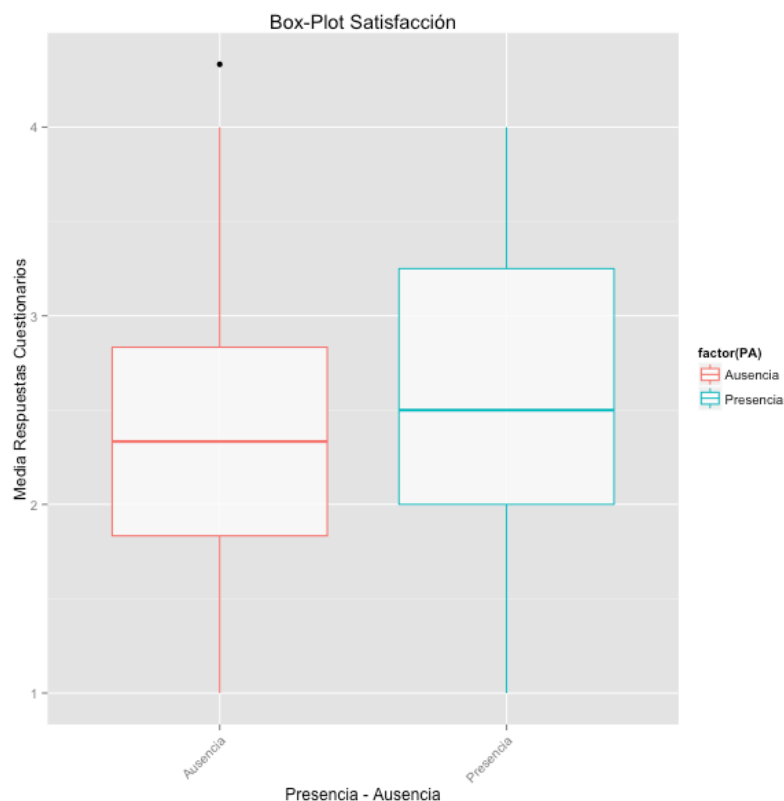


Figura 15: Diagrama de cajas, datos agrupados, para la variable Satisfacción en el experimento piloto

B.4. ANOVA

Para este análisis y de acuerdo a [Field et al., 2012] se ha probado analizar los datos con distintas transformaciones para conseguir algún tipo de resultado en el test ANOVA. A continuación se muestran las tablas de datos obtenidas que contienen tanto el *p-valor* para el test de Shapiro-Wilk que nos permite comprobar la normalidad, el *p-valor* para el test de Levene con el que se verifica la homogeneidad de las poblaciones y el *p-valor* válido para el test ANOVA. En el caso de que no exista ningún factor, o grupo de factores, en el que el *p-valor* sea menor que 0.05 entonces esta casilla aparecerá marcada con ---.

B.4.1. Eficiencia

B.4.1.1. Número de Clicks

En la Tabla 55 se recogen los resultados obtenidos en el análisis ANOVA para la variable número de clicks. Se observa como con tres transformaciones distintas

(ninguna, $\log(X)$ y \sqrt{x}) se obtiene tanto la normalidad de residuos como la homogeneidad de las varianzas (el p -valor es superior a 0.05) pero, sin embargo, en ninguna de las tres situaciones destacadas se obtiene un factor determinante en el test ANOVA realizado. Además, para la transformación $1/X$ no se obtiene normalidad de residuos por lo que no procede el análisis del test ANOVA.

| Transformación | Normalidad | Homogeneidad | ANOVA |
|----------------|------------|--------------|-------|
| Ninguna | 0.0634 | 0.5922 | --- |
| $\log(X)$ | 0.1865 | 0.6299 | --- |
| \sqrt{x} | 0.2985 | 0.6973 | --- |
| $1/X$ | 7.484e-05 | 0.4949 | --- |

Tabla 55: Resultados Eficiencia - N° de Clicks del experimento piloto

B.4.1.2. Tiempo

En la Tabla 56 se recogen los resultados obtenidos en el análisis ANOVA para la variable tiempo. Se observa como con tres transformaciones distintas (ninguna, $\log(X)$ y \sqrt{x}) se obtiene tanto la normalidad de residuos como la homogeneidad de las varianzas (el p -valor es superior a 0.05) pero, sin embargo, en ninguna de las tres situaciones destacadas se obtiene un factor determinante en el test ANOVA realizado. Además, para la transformación $1/X$ no se obtiene normalidad de residuos, aunque no se queda a una distancia tan lejos como en la misma situación de la variable número de clicks, por lo que no procede el análisis del test ANOVA.

| Transformación | Normalidad | Homogeneidad | ANOVA |
|----------------|------------|--------------|-------|
| Ninguna | 0.9713 | 0.5483 | --- |
| $\log(X)$ | 0.5123 | 0.6937 | --- |
| \sqrt{x} | 0.8724 | 0.06312 | --- |
| $1/X$ | 0.0191 | 0.747 | --- |

Tabla 56: Resultados Eficiencia - Tiempo del experimento piloto

B.4.2. Satisfacción

En la Tabla 57 se recogen los resultados obtenidos en el análisis ANOVA para la variable satisfacción. Se observa como con tres transformaciones distintas (ninguna, $\log(X)$ y \sqrt{x}) se obtiene tanto la normalidad de residuos como la homogeneidad de las varianzas (el p -valor es superior a 0.05) pero, sin embargo, en ninguna de las tres situaciones destacadas se obtiene un factor determinante en el test ANOVA realizado. Además, para la transformación $1/X$ no se obtiene normalidad de residuos por lo que no procede el análisis del test ANOVA.

| Transformación | Normalidad | Homogeneidad | ANOVA |
|----------------|------------|--------------|-------|
| Ninguna | 0.694 | 0.1119 | --- |
| log (X) | 0.1371 | 0.1371 | --- |
| sqrt (X) | 0.6536 | 0.1221 | --- |
| 1/X | 0.000148 | 0.1697 | --- |

Tabla 57: Resultados Satisfacción del experimento piloto

B.5. Discusión e Interpretación de Resultados

Se comprueba como en ninguno de los casos el test ANOVA refleja diferencias estadísticas significativas. Es decir, en este caso ninguno de los factores del test (Presencia/Ausencia – Orden – Grupo) es determinante. Por esta razón, en todos los casos habría que aceptar la hipótesis nula y concluir que la presencia de los distintos mecanismos de usabilidad no supone una mejora en la eficiencia y en la satisfacción de los usuarios al realizar las distintas tareas sobre la aplicación *QuickStore*.

Además de los comentarios de los participantes y de un análisis profundo post experimento se desprende la siguiente información:

1. La ausencia del MU-PRE condiciona, sesga, completamente la ejecución de las otras dos tareas aún cuando se utilicen trucos propios del navegador tales como *Zoom*, *FontSizer* o *Traductores*. Esto parece ser un problema:
 - a. Para aquellos sujetos que no pueden hacer las tareas siguientes a PRE.
 - b. Para el propio análisis, ya que la ausencia del PRE influye en las tareas subsecuentes. Es decir, se debería tratar como un ANOVA anidado o sería necesario agrupar todas las tareas y dar una conclusión única.
2. Las evaluaciones, preguntas de los cuestionarios post tarea, se confunden e incluso se mezclan.
3. Sólo el 25% del total de los sujetos era aprovechable para el análisis. Se ha cometido un error conceptual, ya que para el análisis se deben considerar el 50% de los sujetos con el mecanismo presente frente al 50% de los sujetos con el mecanismo inactivo.

Debido a estas conclusiones, se realizan distintas reuniones del equipo de investigación que derivan en el cambio de diseño para el experimento verdadero final presentado en el capítulo 3.

Anexo C. Diseño e Implementación de las Tareas

En este anexo se explica el diseño y la implementación para cada uno de los mecanismos de usabilidad que han sido evaluados durante el experimento. Se describen los escenarios de aplicación y la implementación de cada mecanismo en la funcionalidad seleccionada para la aplicación *QuickStore*.

C.1. Aplicación *QuickStore*

C.1.1. Descripción de la Aplicación

La aplicación seleccionada para la implementación de los mecanismos de usabilidad (ABO, PRO, PRE) es una tienda en línea que gestiona la comercialización de artículos. Permite a los usuarios buscar un determinado artículo, gestionar un carrito de compra, pagar, etc. La tienda cuenta con varias categorías personalizables donde se organizan los diferentes artículos ofrecidos en la tienda. Cuenta con un apartado de informaciones con respecto a las garantías (seguridad, devoluciones y privacidad). Además, brinda una opción de configuración donde los usuarios pueden personalizar la tienda de acuerdo a sus preferencias.

C.1.2. Arquitectura

El desarrollo de la aplicación se basa en la arquitectura MVC (Model-View-Controller) de Java. Una de las principales ventajas de la arquitectura MVC es que reduce el esfuerzo de programación mediante una separación entre los componentes: modelo, vista y controlador. Esta separación posibilita el trabajo en equipo haciendo que las piezas de un sistema puedan ser construidas por separado y luego unir las en tiempo de ejecución. Si uno de los componentes presenta un mal funcionamiento, puede ser reemplazado sin afectar a los demás. En la Figura 16 se observa el patrón arquitectónico MVC para la aplicación de la tienda en línea.

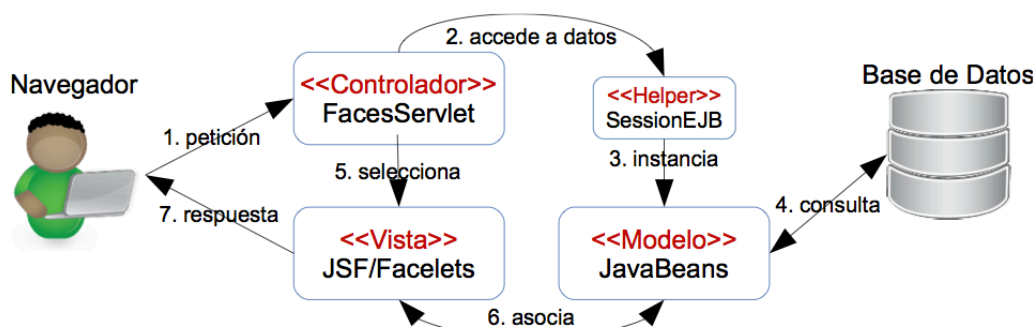


Figura 16: Arquitectura MVC de Java

C.2. Mecanismo de Usabilidad Abortar Operación

Para la implementación del mecanismo de usabilidad Abortar Operación se ha utilizado el patrón de programación propuesto por [Rodríguez, 2014]. En la Figura 17 se muestra el árbol de escenarios de aplicación del ABO en un formulario de entrada de datos. Este mecanismo ha sido implementado en la funcionalidad de Carrito de Compra.

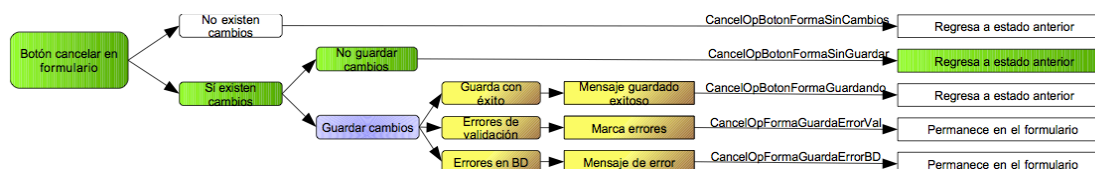


Figura 17: Árbol de escenarios de aplicación del MU-ABO

En la Tabla 58 se describen las precondiciones, la tarea solicitada al usuario y la secuencia de acciones esperadas atendiendo la presencia o ausencia del mecanismo ABO.

| | |
|---|--|
| <p>Precondiciones</p> <ul style="list-style-type: none"> El usuario debe acceder (autologin) a la tienda con una cuenta/contraseña generada por sistema El carrito del usuario ya contendrá al menos 1 artículo añadido | |
| <p>Descripción de la tarea:</p> <p>Imagina que en el día de ayer visitaste la tienda y ya tienes ítems o artículos agregados a tu carrito. Accede entonces a tu carrito de compra y de la lista de artículos ya agregados, modifica la cantidad del primer ítem/artículo ingresando la cantidad que desees, luego ingresa el código promocional “2015” en la casilla especificada y valida dicho código para tener el precio final.</p> <p>Ahora te das cuenta de que el precio final supera tus expectativas y te arrepientes, por lo que decides comprar en otro momento y anulas las acciones para dejar sin efecto los cambios que acabas de realizar en tu carrito.</p> | <p>Secuencia de acciones esperada:</p> <ol style="list-style-type: none"> El usuario hace click en el carrito Incrementa la cantidad del primer ítem Introduce el código 2015 en el campo “Código Promocional” El usuario desea cancelar: <ol style="list-style-type: none"> Con MU: El usuario hace click en el botón Cancelar y no guarda los cambios. Sin MU: Botón cancelar ausente. El usuario intentará deshacer los cambios manualmente, cerrar la aplicación, seleccionar otra opción, refrescar la página, no hacer nada, etc. |

Tabla 58: Contexto, tareas y secuencia de acciones del MU-ABO

La interfaz implementada del Carrito de Compra para la evaluación del MU-ABO se visualiza en la Figura 18. Se trata de un formulario de entrada de datos en la cual se implementa el árbol de escenarios de aplicación del MU-ABO mediante el botón Cancelar.

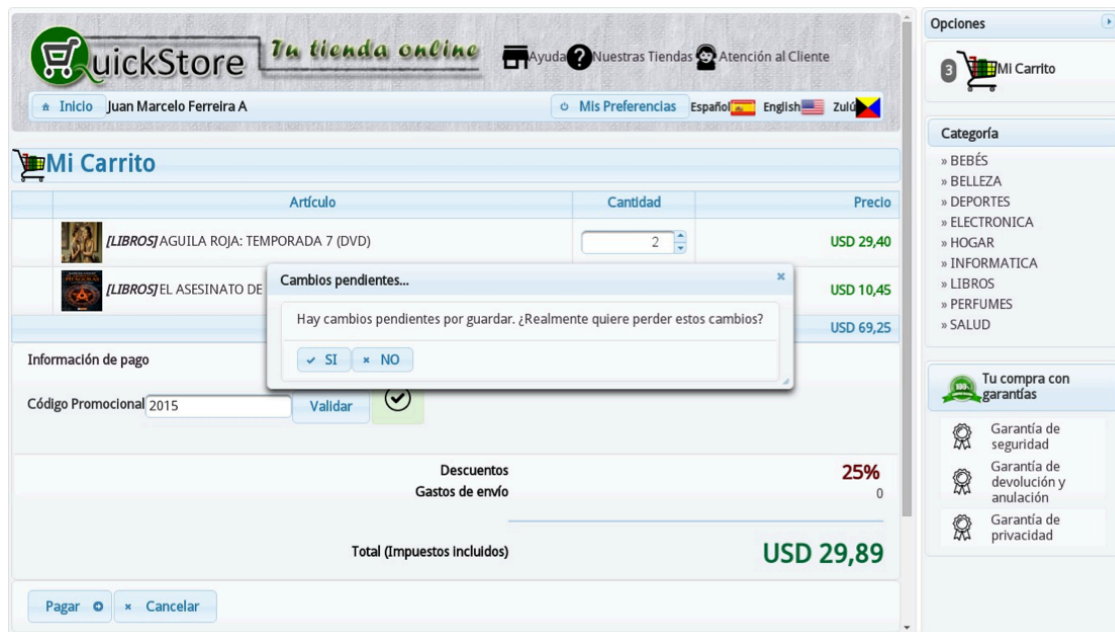


Figura 18: Interfaz de tarea para el MU-ABO

Considerando el árbol de escenarios de aplicación se espera que ante cualquier cambio en el formulario exista una alerta de confirmación de las acciones. Esta alerta debe declarar el cambio de estado en el formulario y qué acciones se realizarán con ese cambio. En la Figura 18 se visualiza el mensaje de confirmación solicitando al usuario qué acciones tomar cuando el MU-ABO está presente.

Para recolectar las métricas de satisfacción se ha diseñado el cuestionario de satisfacción pos-tarea. Dicho cuestionario intenta recolectar aspectos con los que el usuario está conforme o disconforme.

| Sentencias | Totalmente en desacuerdo | En desacuerdo | Neutral | De acuerdo | Totalmente de acuerdo |
|--|--------------------------|---------------|---------|------------|-----------------------|
| 1. He logrado hacer la cancelación | | | | | |
| 2. Me sentí cómodo haciendo la cancelación de las modificaciones que no deseo en mi carrito de compra. | | | | | |
| 3. Siento que el sistema me ayuda porque me provee alerta de los cambios hechos en mi carrito de compra y me solicita confirmación para proceder | | | | | |

Tabla 59: Diseño del cuestionario de satisfacción post-tarea del MU-ABO

C.3. Mecanismo de Usabilidad Retroalimentación de Progreso

Análogamente al MU-ABO, se ha utilizado el árbol de escenarios propuesto por [Rodríguez, 2014] para el MU-PRO. Dicho árbol de escenarios se visualiza en la

Figura 19. Este mecanismo ha sido implementado en la funcionalidad de Búsqueda de Artículos.

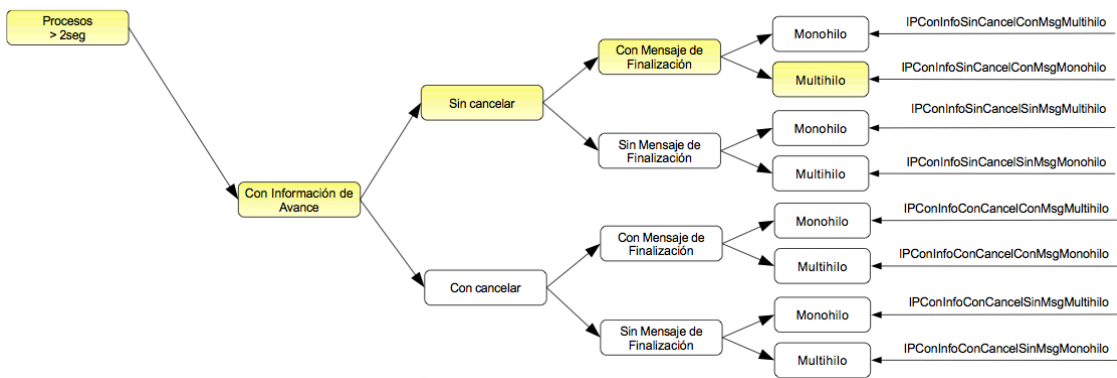


Figura 19: Árbol de escenarios de aplicación del MU-PRO

En la Tabla 60 se describen las precondiciones, la tarea solicitada al usuario y la secuencia de acciones esperadas atendiendo la presencia o ausencia del mecanismo PRO.

| | |
|--|---|
| Precondiciones: Sin precondiciones | |
| Descripción de la tarea: Imagina que un amigo te ha recomendado visitar la tienda porque estás interesado en adquirir el libro titulado “LA CHICA DEL TREN”. Entonces realiza una búsqueda de este artículo y agrégalo a tu carrito de compra. | Secuencia de acciones esperada: <ol style="list-style-type: none">El usuario introduce “LA CHICA DEL TREN” en el campo de búsqueda.El usuario hace <i>click</i> en el botón buscar.<ol style="list-style-type: none">Con MU: El sistema muestra una barra de progreso con información de avance mientras se ejecuta la búsqueda.Sin MU: El sistema no muestra nada (permanece congelada) mientras se ejecuta la búsqueda.El usuario hace <i>click</i> en el botón agregar del artículo visualizado. |

Tabla 60: Contexto, tareas y secuencia de acciones del MU-PRO

La interfaz implementada de la Búsqueda de Artículos para la evaluación del MU-PRO se visualiza en la Figura 20. Se trata de un proceso de búsqueda ejecutándose en segundo plano, mientras se visualiza información de avance de la consulta, que no permite cancelación y emitiendo un mensaje al final indicando la cantidad de registros recuperados. El camino recorrido de esta tarea se muestra resaltado en el árbol de la Figura 19.

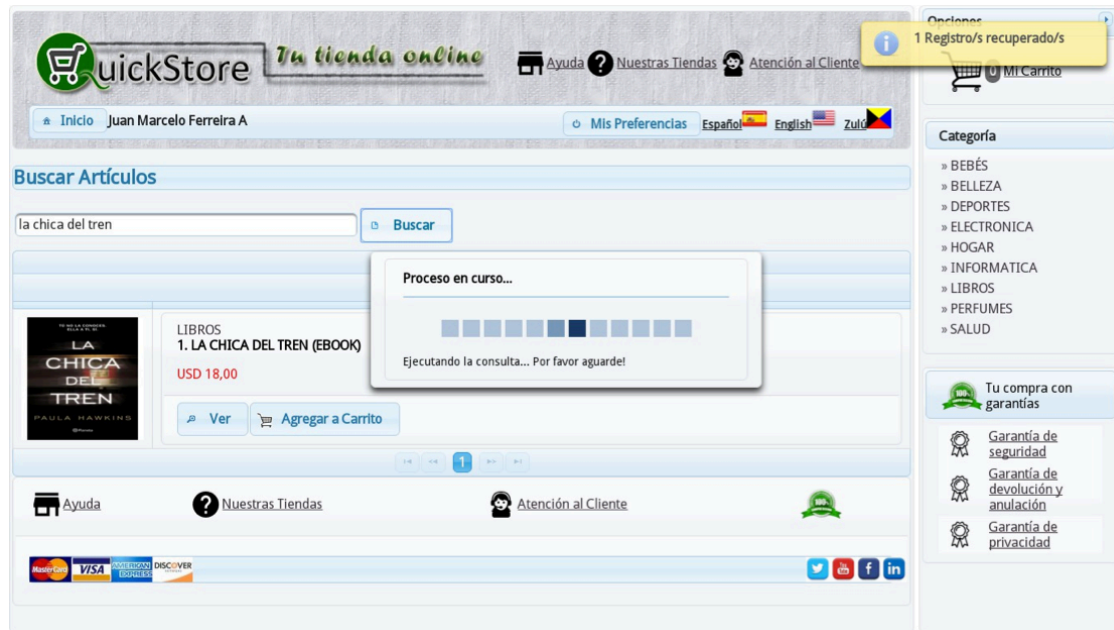


Figura 20: Interfaz de tarea para el MU-PRO

El cuestionario post tarea usado para el MU-PRO se muestra en la Tabla 61.

| Sentencias | Totalmente en desacuerdo | En desacuerdo | Neutral | De acuerdo | Totalmente de acuerdo |
|---|--------------------------|---------------|---------|------------|-----------------------|
| 1. He logrado buscar el artículo deseado. | | | | | |
| 2. He tenido algún tipo de retroalimentación mientras estaba realizando la búsqueda. Es decir, en todo momento el sistema me informaba de lo que estaba haciendo. | | | | | |
| 3. Me he sentido orientado/ayudado por el sistema cuando hice la búsqueda del artículo. | | | | | |

Tabla 61: Diseño del cuestionario de satisfacción post-tarea del MU-PRO

C.4. Mecanismo de Usabilidad Preferencias

El MU-PRE es una funcionalidad tratada independiente de cualquier funcionalidad del sistema. Sin embargo, sus efectos se propagan por la aplicación web en su totalidad. En [Rodríguez, 2010] se aborda este mecanismo como un requisito funcional detallando sus casos de uso, modelo conceptual y diagramas de secuencias. En este sentido, las configuraciones propuestas son: lenguaje, familia y tamaño de fuente, esquemas de color, iconos y sonido. Para el diseño experimental se ha considerado evaluar el tamaño y familia de fuente.

En la Tabla 62 se describen las precondiciones, tareas y secuencia de acciones esperada por parte del usuario. Téngase en cuenta que ante la ausencia del MU-PRE es probable que el usuario no intente estrategias fuera del navegador quedando completamente imposibilitado en hacer la tarea.

| | |
|---|--|
| <p>Precondiciones</p> <p>El usuario debe acceder (autologin) a la tienda con una cuenta/contraseña generada por sistema</p> | |
| <p>Descripción de la tarea:</p> <p>Revisa la apariencia de la Tienda Virtual y comprueba que estás cómodo con sus características: tamaño de letra, tipo de letra... Si necesitas cambiar algo, ahora es el momento. Procede, por favor, a hacerlo.</p> | <p>Secuencia de acciones esperada:</p> <ol style="list-style-type: none"> Con MU: <ol style="list-style-type: none"> El usuario ingresa a la opción de Mis Preferencias. El usuario modifica uno o más parámetros (tamaño y tipo de fuente) que considera apropiados para su uso. El usuario hace click en guardar para que el sistema aplique las configuraciones visuales. Sin MU: <ol style="list-style-type: none"> El usuario podría proceder de distintas formas utilizando estrategias de Zoom, Copy&Paste o bien, sentirse imposibilitado de cambiar las configuraciones visuales. |

Tabla 62: Contexto, tareas y secuencia de acciones del MU-PRE

La interfaz implementada para el cambio de las configuraciones visuales se visualiza en la Figura 21. Si bien se ha implementado el cambio del idioma, la tarea que el usuario debe hacer declara solo el cambio del tamaño y tipo de fuente. En la Tabla 63 se puede comprobar el cuestionario post tarea para medir la satisfacción.

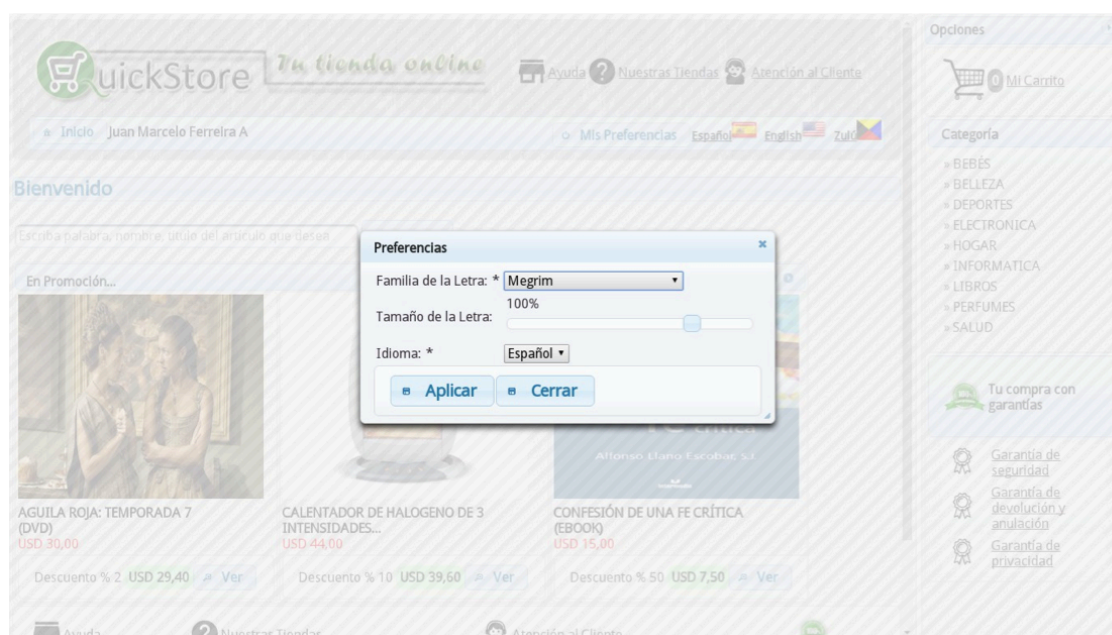


Figura 21: Interfaz de tarea para el MU-PRE

| Sentencias | Totalmente en desacuerdo | En desacuerdo | Neutral | De acuerdo | Totalmente de acuerdo |
|---|--------------------------|---------------|---------|------------|-----------------------|
| 1. He logrado poner la interfaz del sistema de acuerdo a mis preferencias (tamaño y tipo de letra). | | | | | |
| 2. He logrado encontrar la información de los plazos para la devolución de artículos. | | | | | |
| 3. Pude encontrar, de manera sencilla, la información del plazo para la devolución con las configuraciones visuales de la tienda. | | | | | |
| 4. Las configuraciones visuales de la tienda me han ayudado a localizar más rápidamente la información del plazo. | | | | | |

Tabla 63: Diseño del cuestionario de satisfacción post-tarea del MU-PRE

Anexo D. Diagramas de Caja del Experimento Verdadero

D.1. Eficiencia

D.1.1. Número de Clicks

D.1.1.1. Preferencias

En la Figura 22 está representado el diagrama de cajas para el mecanismo Preferencias considerando los factores Presencia/Ausencia y Grupo para la variable respuesta número de clicks. No se observa una diferencia significativa entre los cuatro grupos aunque los grupos que más se parecen, sorprendentemente son el C-Ausencia y A-Presencia con medianas parecidas, Q1 muy parecidos y tres outliers en cada caso.

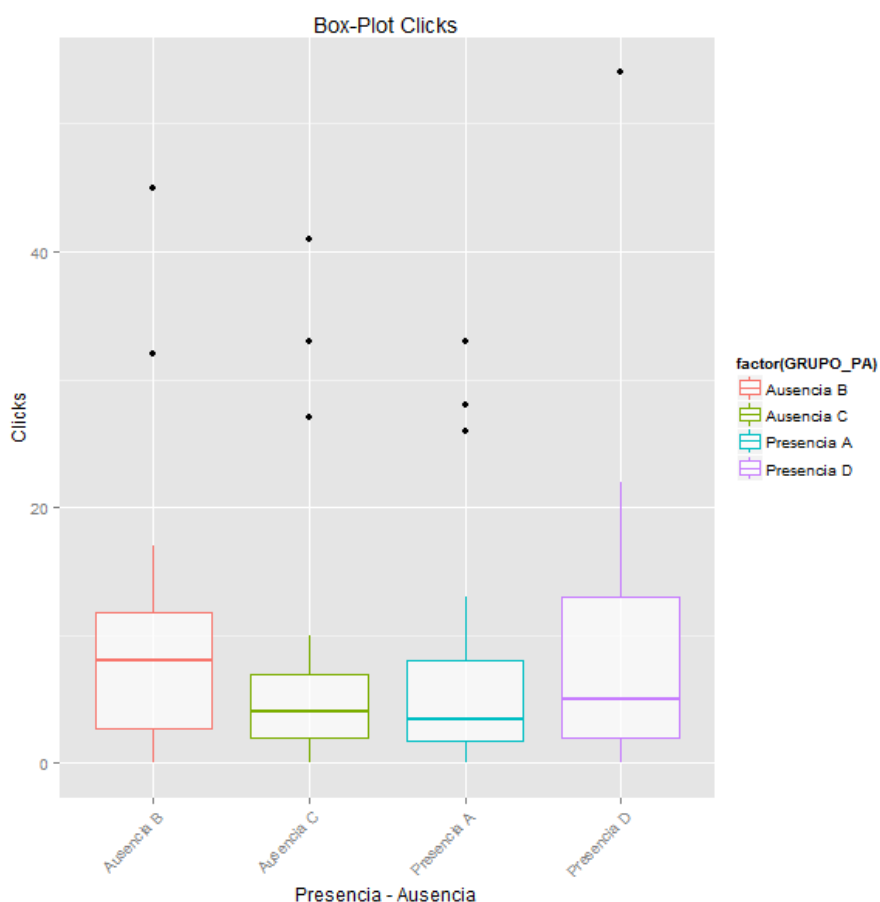


Figura 22: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRE y N° de Clicks

En la Figura 23 aparece el diagrama de cajas del mecanismo Preferencias para los factores Presencia/Ausencia y Orden y la variable respuesta número de clicks. Se observa que los órdenes 1 y 2 tienen un comportamiento muy similar tanto en el caso de la presencia como de la ausencia del mecanismo. A su vez, este comportamiento es muy diferente al que se observa en el orden 5, sobre todo en ausencia del mecanismo.

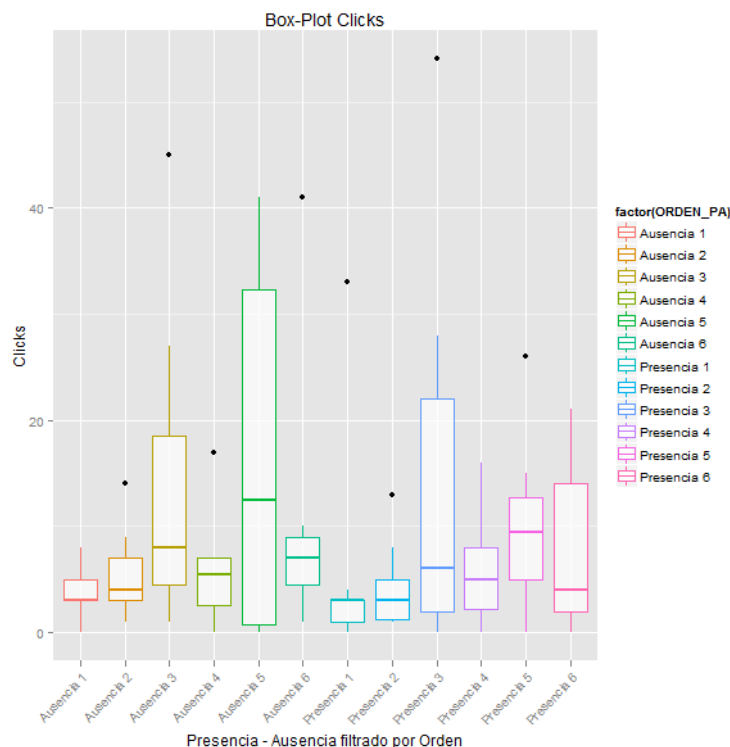


Figura 23: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRE y N° de Clicks

D.1.1.2. Retroalimentación de Progreso

En la Figura 24 aparece el diagrama de cajas del mecanismo Retroalimentación de Progreso para los factores Presencia/Ausencia y Grupo y la variable respuesta número de clicks. Se observa que el tamaño de las cajas es parecido en el caso de los grupos que tienen el mecanismo presente mientras que, en el caso de los grupos que no disponían del mecanismo las dos cajas son más dispares. Además, los grupos A y C tienen dos outliers mientras que en las otras dos submuestras no aparecen valores atípicos.

En la Figura 25 aparece el diagrama de cajas del mecanismo Retroalimentación de Progreso para los factores Presencia/Ausencia y Orden y la variable respuesta número de clicks. En este diagrama es muy difícil observar relaciones directas entre los órdenes con presencia o ausencia del mecanismo si bien destacan los órdenes 2 y 3 en ausencia como las submuestras que toman unos valores más altos y tienen además una caja más alargada.

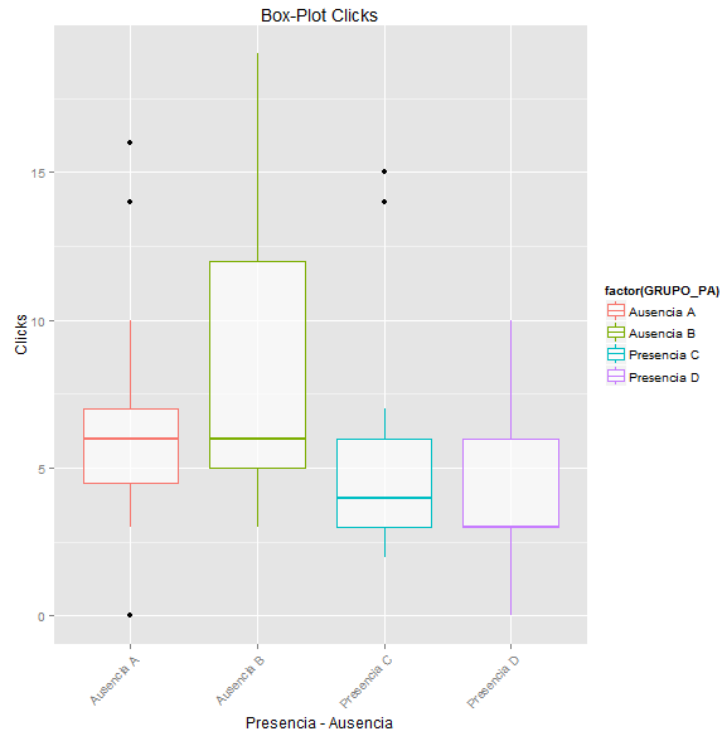


Figura 24: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRO y N° de Clicks

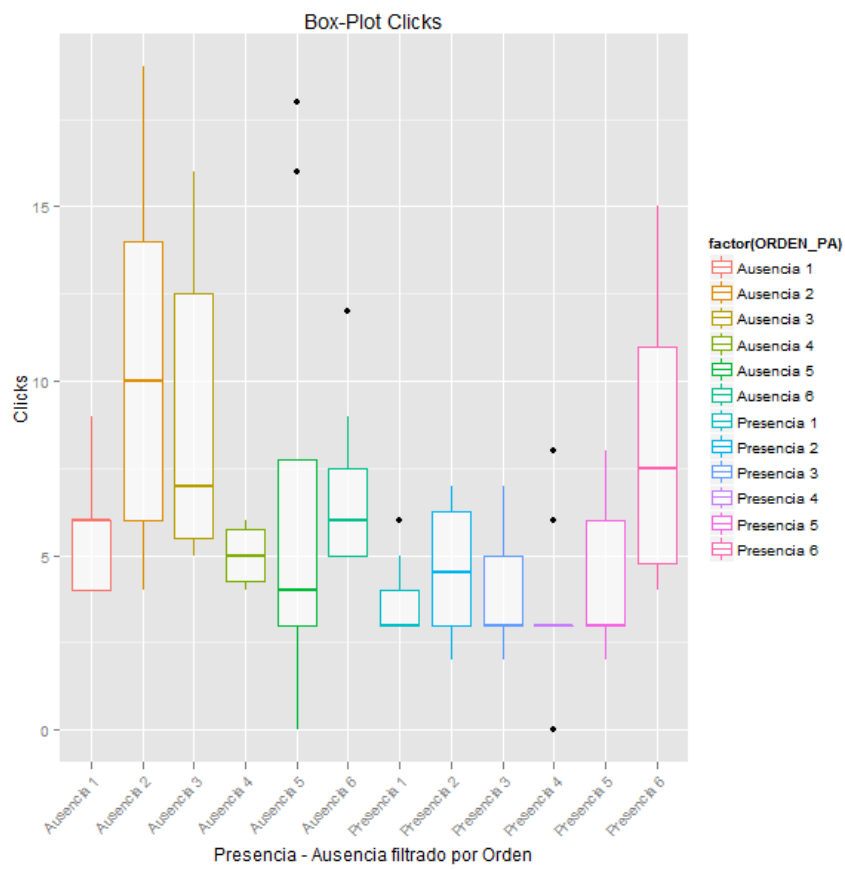


Figura 25: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRO y N° de Clicks

D.1.1.3. Abortar Operación

En la Figura 26 aparece el diagrama de cajas del mecanismo Abortar Operación para los factores Presencia/Ausencia y Grupo y la variable respuesta número de clicks. Como sucedía en el caso de la Retroalimentación de Progreso se observa como parece que el grupo no es determinante sino que lo que hace que el aspecto de la caja varíe es la presencia o ausencia del mecanismo.

En la Figura 27 aparece el diagrama de cajas del mecanismo Abortar Operación para los factores Presencia/Ausencia y Orden y la variable respuesta número de clicks. En este diagrama parece que los órdenes se comportan igual independientemente de la presencia o ausencia del mecanismo. Por ejemplo las cajas para el orden 4 son muy similares tanto en presencia como en ausencia del mecanismo.

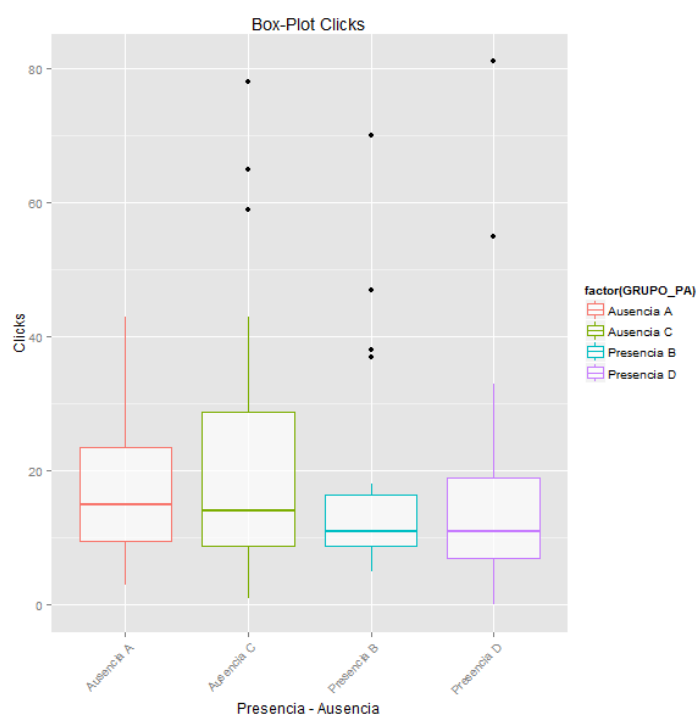


Figura 26: Diagrama de cajas Presencia/Ausencia y Grupo para MU-ABO y N° de Clicks

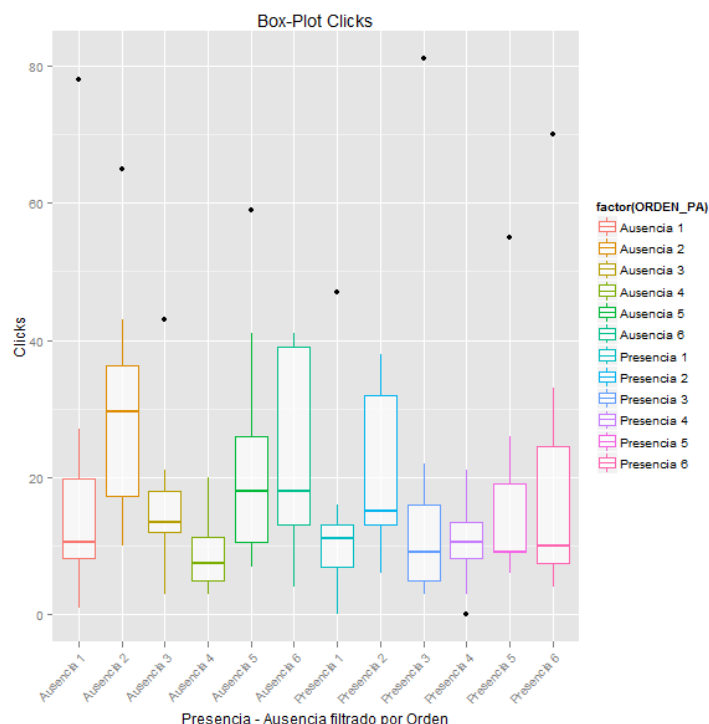


Figura 27: Diagrama de cajas Presencia/Ausencia y Orden para MU-ABO y N° de Clicks

D.1.2. Tiempo

D.1.2.1. Preferencias

En la Figura 28 aparece el diagrama de cajas del mecanismo Preferencias para los factores Presencia/Ausencia y Grupo y la variable respuesta tiempo. Se observa como las cuatro cajas son muy similares destacando únicamente que en Ausencia-B hay dos outliers mientras que en las demás submuestras sólo hay uno.

En la Figura 29 aparece el diagrama de cajas del mecanismo Preferencias para los factores Presencia/Ausencia y Orden y la variable respuesta tiempo. En este diagrama parece que los órdenes se comportan igual independientemente de la presencia o ausencia del mecanismo salvando el caso del orden 6 que presenta una forma más alargada en el caso de la presencia que en el caso de la ausencia. Destacar que en toda la muestra sólo hay tres outliers uno en cada una de las siguientes submuestras: Ausencia-3, Presencia-3, Ausencia-6.

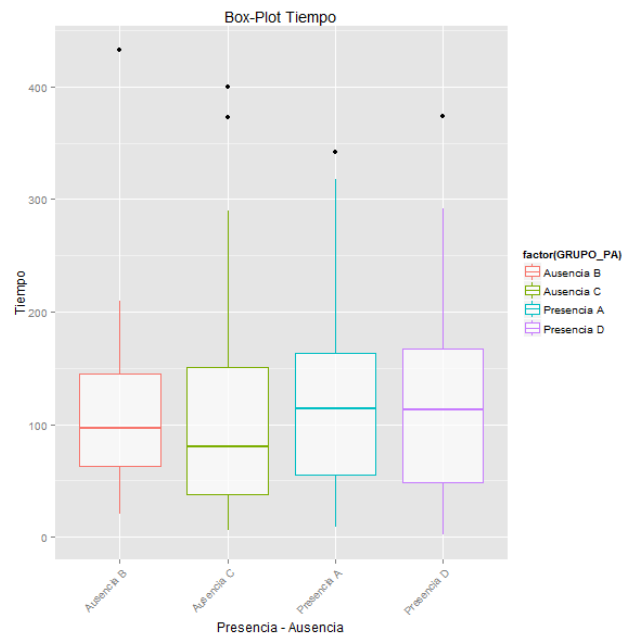


Figura 28: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRE y Tiempo

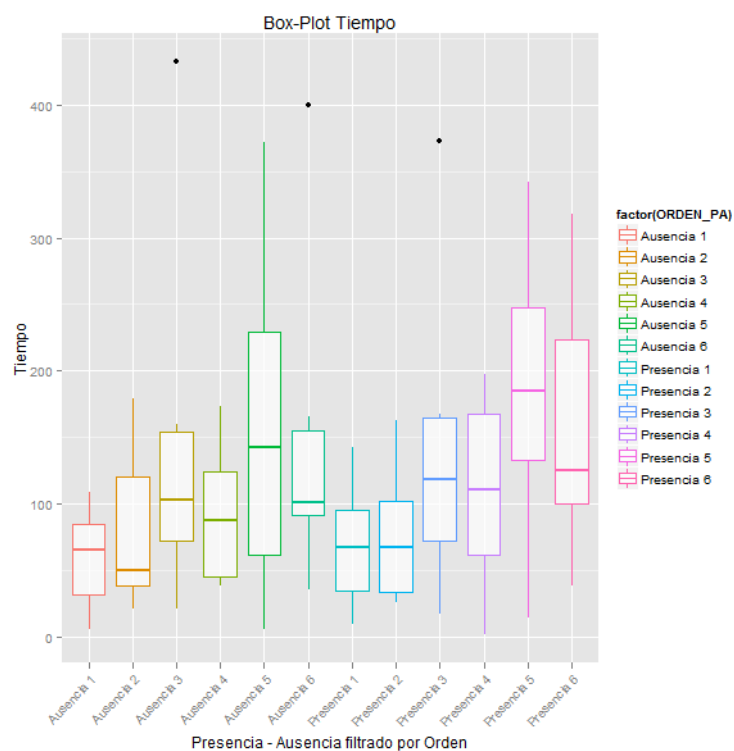


Figura 29: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRE y Tiempo

D.1.2.2. Retroalimentación de Progreso

En la Figura 30 aparece el diagrama de cajas del mecanismo Retroalimentación de Progreso para los factores Presencia/Ausencia y Grupo y la variable respuesta tiempo. Se observa que las cajas más parecidas son Ausencia-A y Presencia-C salvo que la primera presenta unos bigotes mucho más pronunciados. Sin embargo, las otras dos restantes no se parecen demasiado.

En la Figura 31 aparece el diagrama de cajas del mecanismo Retroalimentación de Progreso para los factores Presencia/Ausencia y Orden y la variable respuesta tiempo. Se observa como el comportamiento es muy diferente entre muchos de estos órdenes. Se puede decir que los órdenes 1, 5 y 6 se comportan igual independientemente del factor Presencia/Ausencia mientras que los tres restantes difieren bastante dependiendo de si el mecanismo está activo o no.

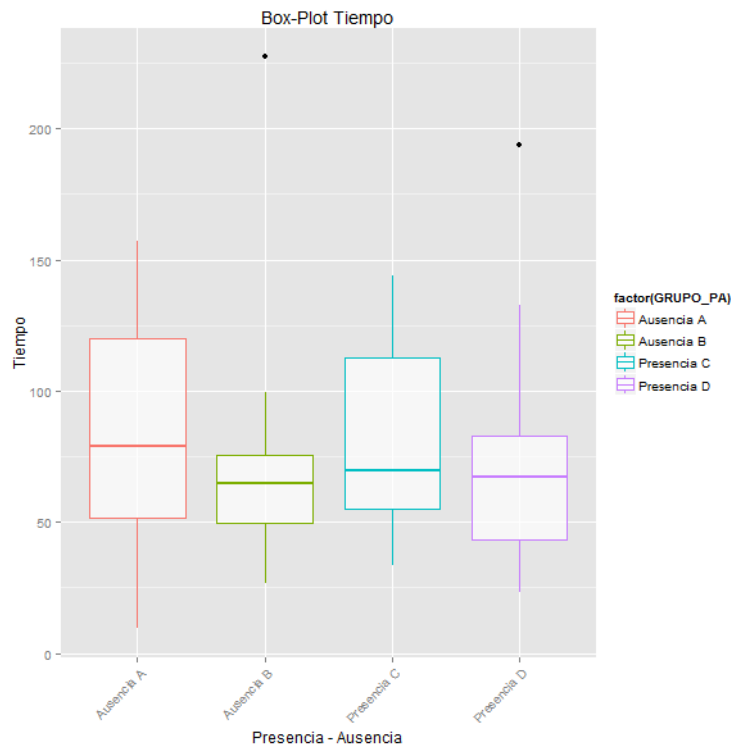


Figura 30: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRO y Tiempo

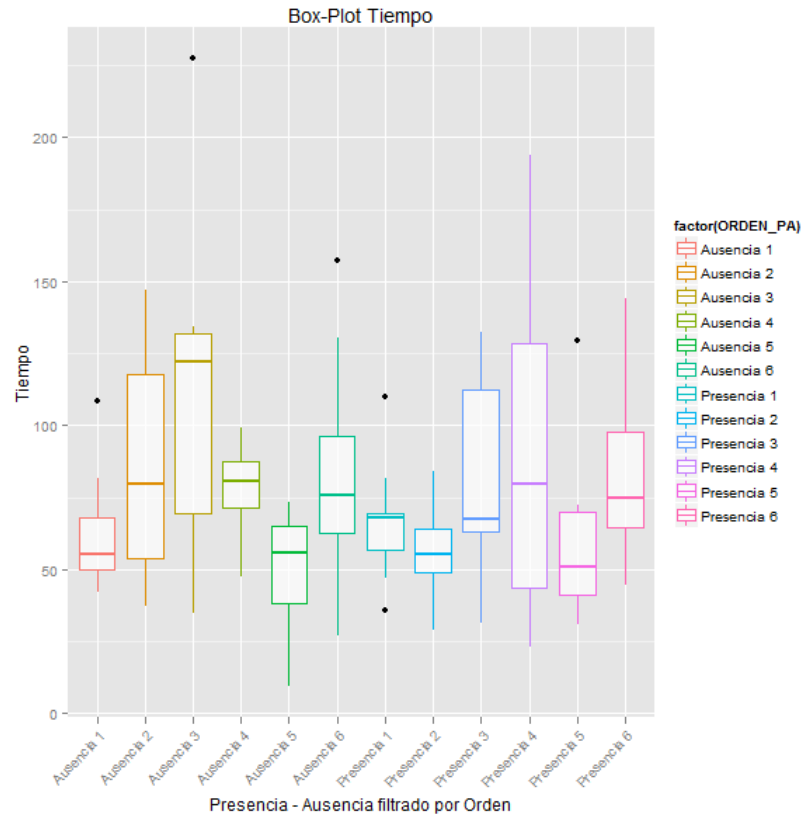


Figura 31: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRO y Tiempo

D.1.2.3. Abortar Operación

En la Figura 32 aparece el diagrama de cajas del mecanismo Abortar Operación para los factores Presencia/Ausencia y Grupo y la variable respuesta tiempo. Se observa un comportamiento idéntico en el caso de la ausencia del mecanismo, salvo porque en el grupo B hay un valor atípico y en el grupo A no hay ninguno. En las dos submuestras que disponen del mecanismo activo los comportamientos son más dispares.

En la Figura 33 aparece el diagrama de cajas del mecanismo Abortar Operación para los factores Presencia/Ausencia y Orden y la variable respuesta tiempo. Se observa como la tendencia es que todas las cajas con ausencia del mecanismo están posicionadas un poco más arriba, es decir, los sujetos han necesitado más tiempo, que en el caso de las cajas referidas a sujetos con presencia del mecanismo. No se observa una identificación clara entre los órdenes dependiendo de si el mecanismo está activo o no.

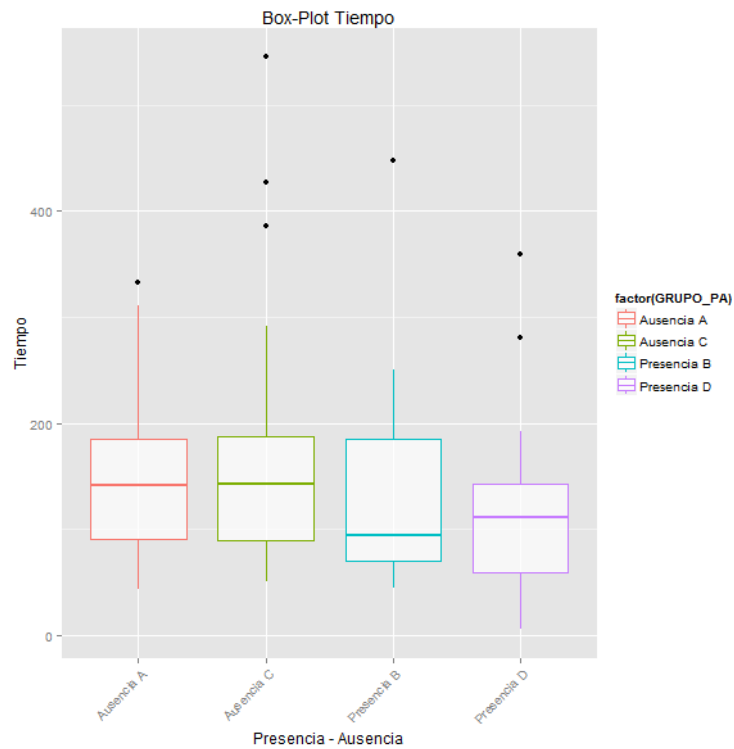


Figura 32: Diagrama de cajas Presencia/Ausencia y Grupo para MU-ABO y Tiempo

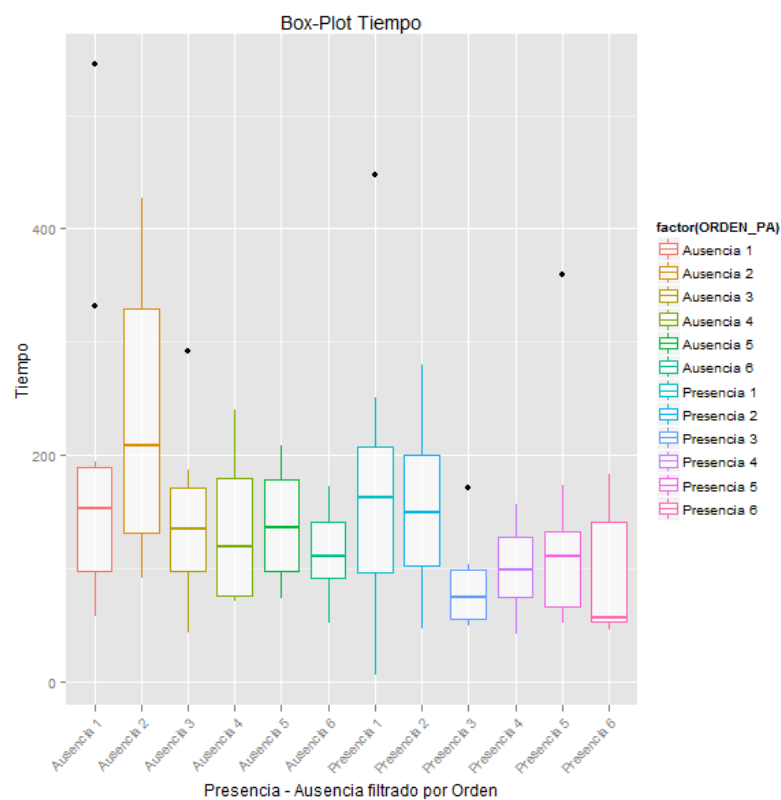


Figura 33: Diagrama de cajas Presencia/Ausencia y Orden para MU-ABO y Tiempo

D.2. Satisfacción

D.2.1. Preferencias

En la Figura 34 aparece el diagrama de cajas del mecanismo Preferencias para los factores Presencia/Ausencia y Grupo y la variable respuesta satisfacción. Se observa como claramente la satisfacción es mucho superior en el caso de la presencia del mecanismo teniendo que el primer cuartil para esta submuestra es superior al tercer cuartil de la muestra restante. Destaca también como el comportamiento es muy similar en el caso de los sujetos que no disponían del mecanismo (grupos B y C) mientras es un poco diferente en el caso de los grupos que tenían el mecanismo presente (grupos A, D).

En la Figura 35 aparece el diagrama de cajas del mecanismo Preferencias para los factores Presencia/Ausencia y Orden y la variable respuesta satisfacción. Al igual que en el diagrama anterior los valores que toma la satisfacción son más altos en el caso de la presencia del mecanismo independientemente del factor orden. Destaca que la satisfacción es nula en las submuestras Ausencia-3 y Ausencia-4.

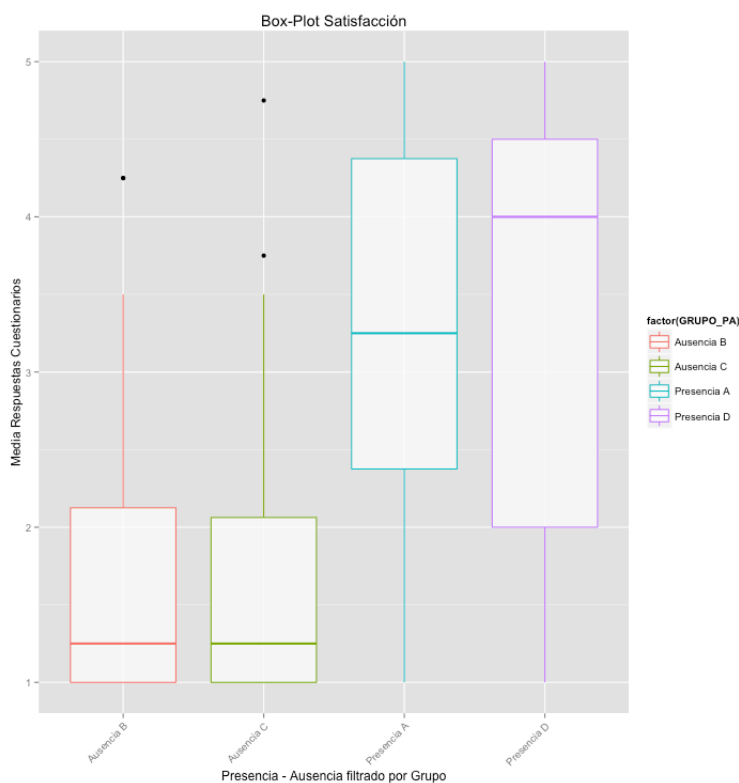


Figura 34: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRE y Satisfacción

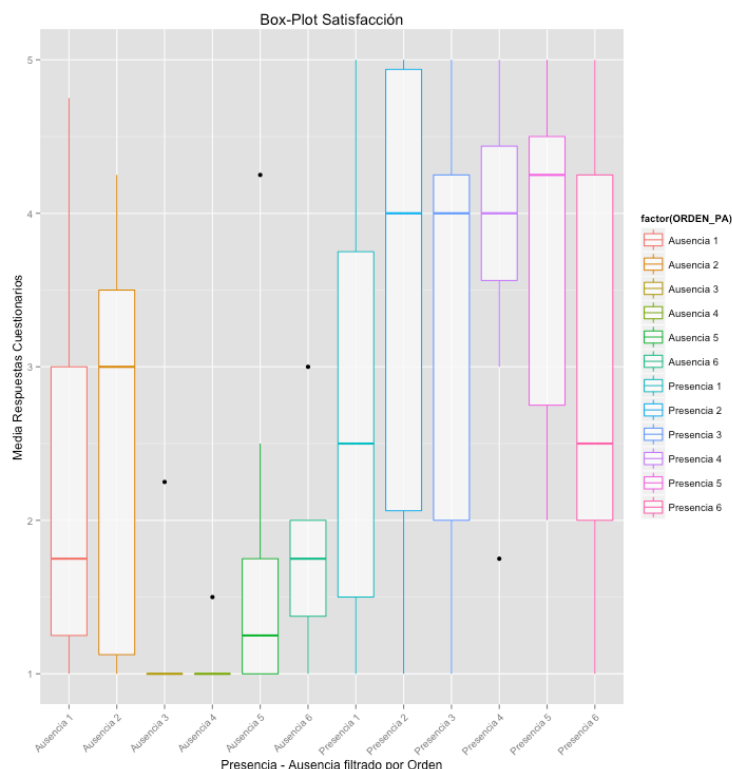


Figura 35: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRE y Satisfacción

D.2.2. Retroalimentación de Progreso

En la Figura 36 aparece el diagrama de cajas del mecanismo Retroalimentación de Progreso para los factores Presencia/Ausencia y Grupo y la variable respuesta satisfacción. Destaca que la satisfacción es alta pues el diagrama, en los cuatro casos, está posicionado entre el valor tres y cinco. Si bien las medianas y primeros cuartiles son idénticos en el caso de la presencia del mecanismo (grupos C y D) éstos valores difieren cuando no se dispone del mecanismo Retroalimentación de Progreso. Destaca también que el bigote inferior del grupo C es más corto que el bigote para el grupo D.

En la Figura 37 aparece el diagrama de cajas del mecanismo Retroalimentación de Progreso para los factores Presencia/Ausencia y Orden y la variable respuesta satisfacción. Al igual que en el diagrama anterior los valores que toma la satisfacción son altos para casi la totalidad de la muestra. Sin embargo, en este caso es mucho más difícil sacar conclusiones acerca de cómo influye el orden de exposición pues no se encuentra una relación directa entre los órdenes cuando el mecanismo está presente o ausente.

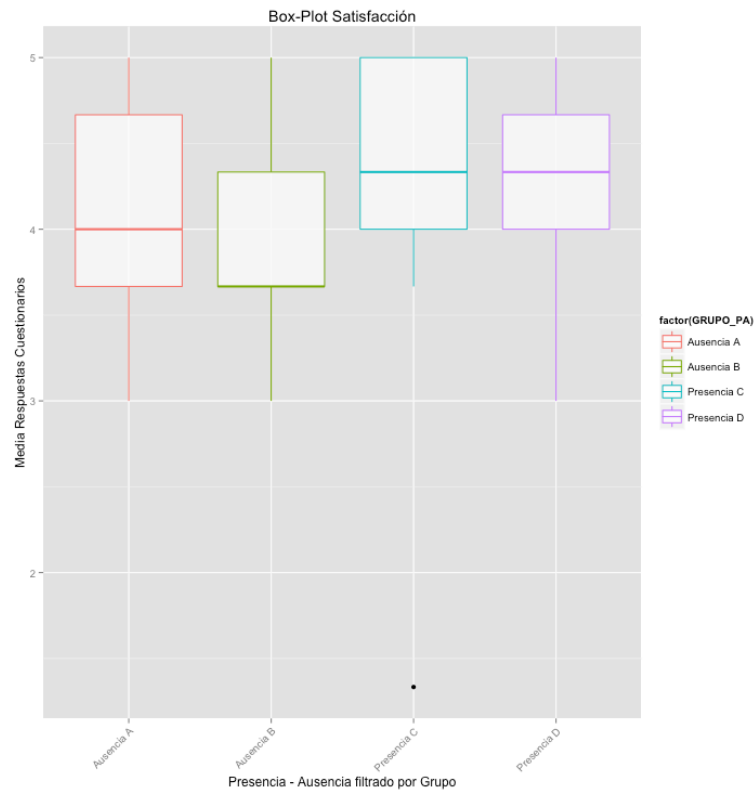


Figura 36: Diagrama de cajas Presencia/Ausencia y Grupo para MU-PRO y Satisfacción

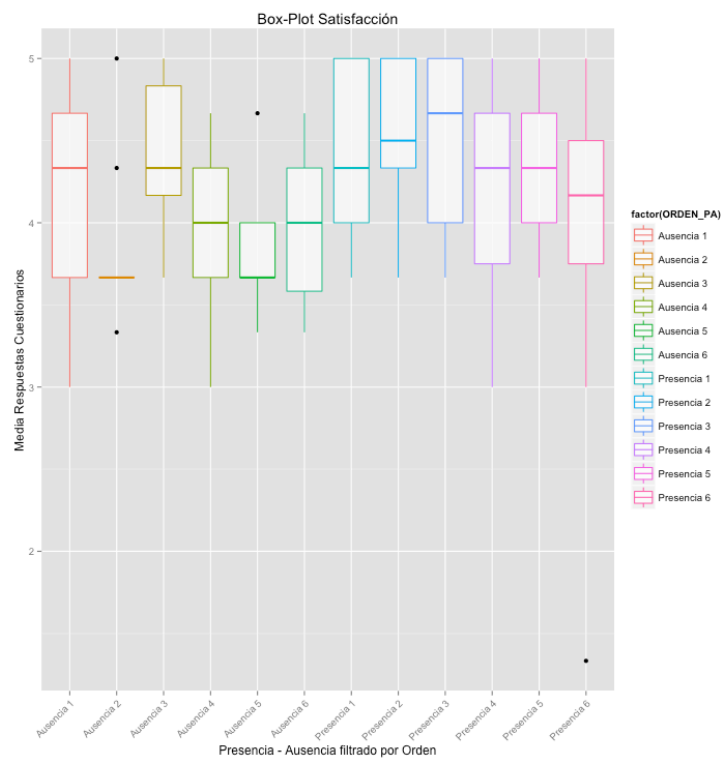


Figura 37: Diagrama de cajas Presencia/Ausencia y Orden para MU-PRO y Satisfacción

D.2.3. Abortar Operación

En la Figura 38 aparece el diagrama de cajas del mecanismo Abortar Operación para los factores Presencia/Ausencia y Grupo y la variable respuesta satisfacción. Se observa como claramente la satisfacción es muy superior cuando el mecanismo está presente. Destaca como el comportamiento entre los grupos que tienen el mecanismo activo (B y D) y los grupos que no tienen el mecanismo activo (A y C) es bastante diferente entre ellos.

En la Figura 39 aparece el diagrama de cajas del mecanismo Abortar Operación para los factores Presencia/Ausencia y Orden y la variable respuesta satisfacción. Al igual que en el diagrama anterior, la tendencia es que las submuestras que tenían el mecanismo presente toman valores muy superiores a las muestras que no disponían del mecanismo independientemente del orden de exposición.

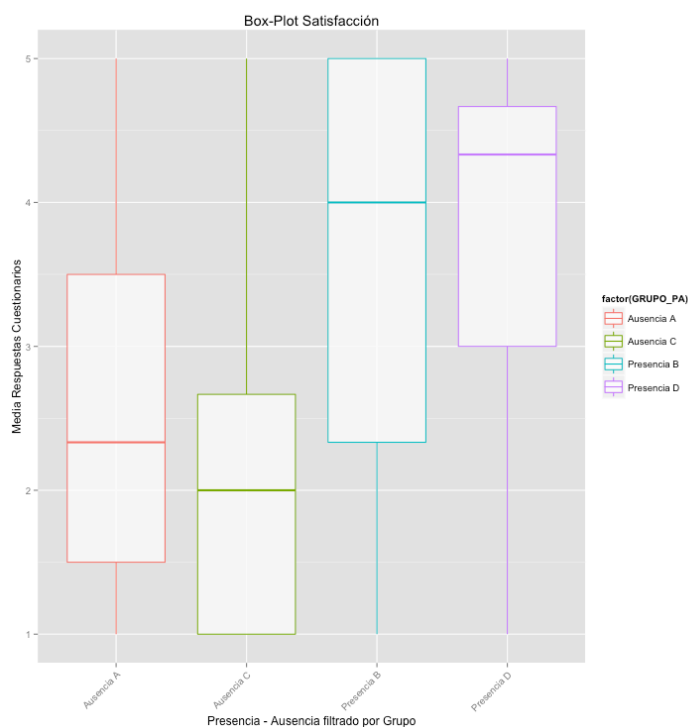


Figura 38: Diagrama de cajas Presencia/Ausencia y Grupo para MU-ABO y Satisfacción

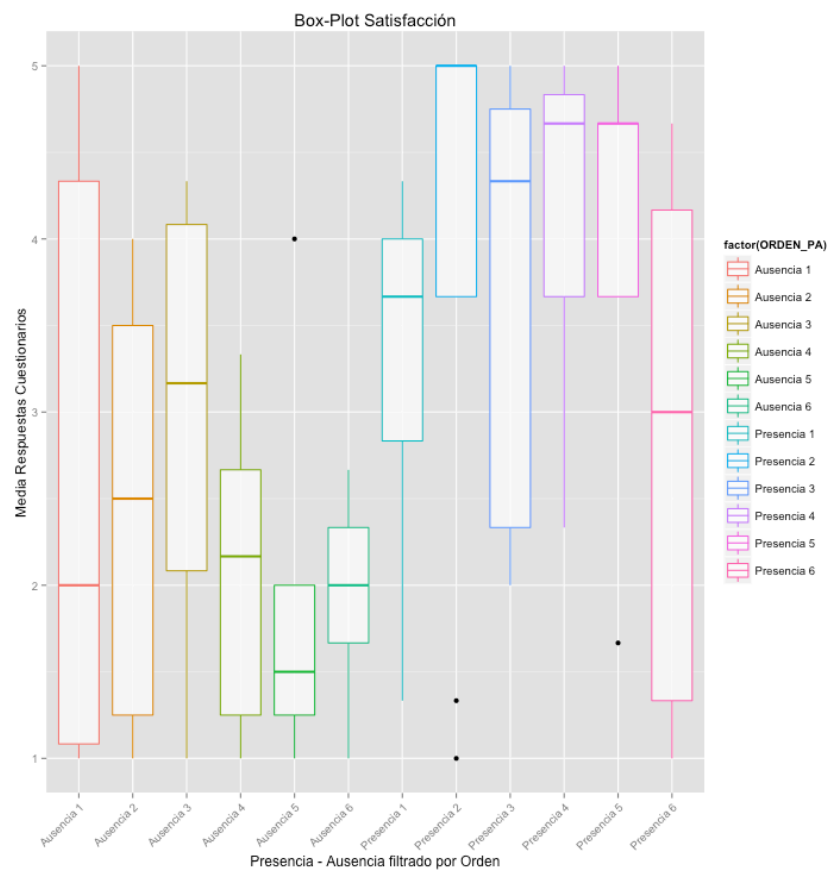


Figura 39: Diagrama de cajas Presencia/Ausencia y Orden para MU-ABO y Satisfacción

Anexo E. Test de Tukey

En las Figuras 40, 41, 42, 43 y 44 aparece la representación gráfica de los diagramas de Tukey utilizados para complementar el test ANOVA cuando éste determina que el factor determinante es el orden.

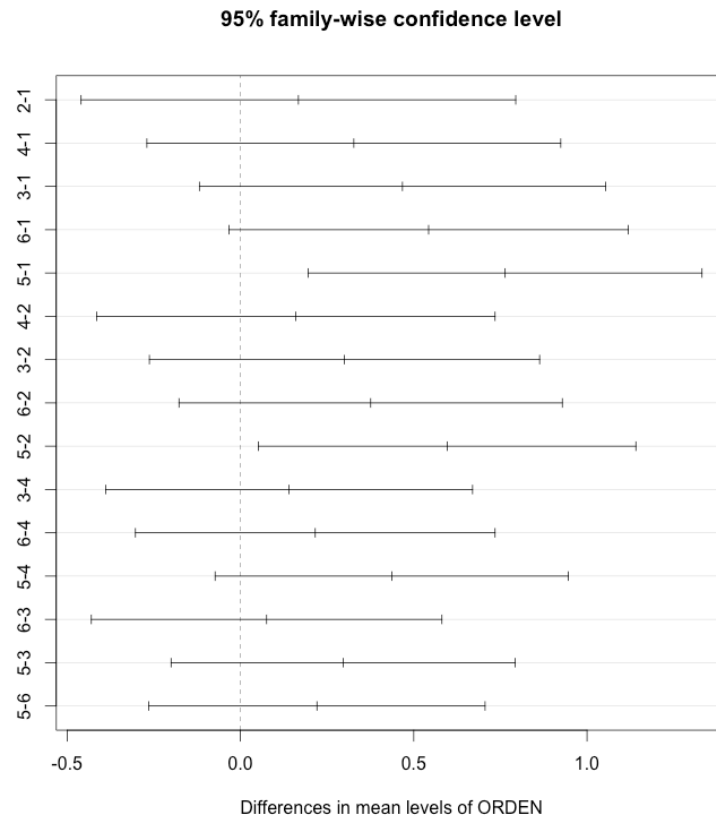


Figura 40: Test de Tukey para el factor Orden y la variable Tiempo MU-PRE

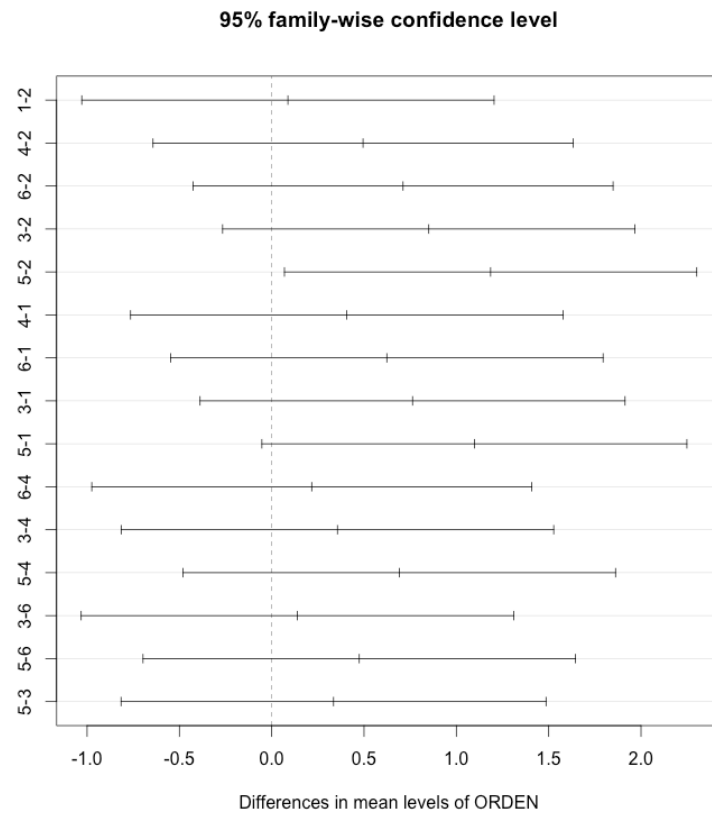


Figura 41: Test de Tukey para el factor Orden y la variable N° de Clicks MU-PRE

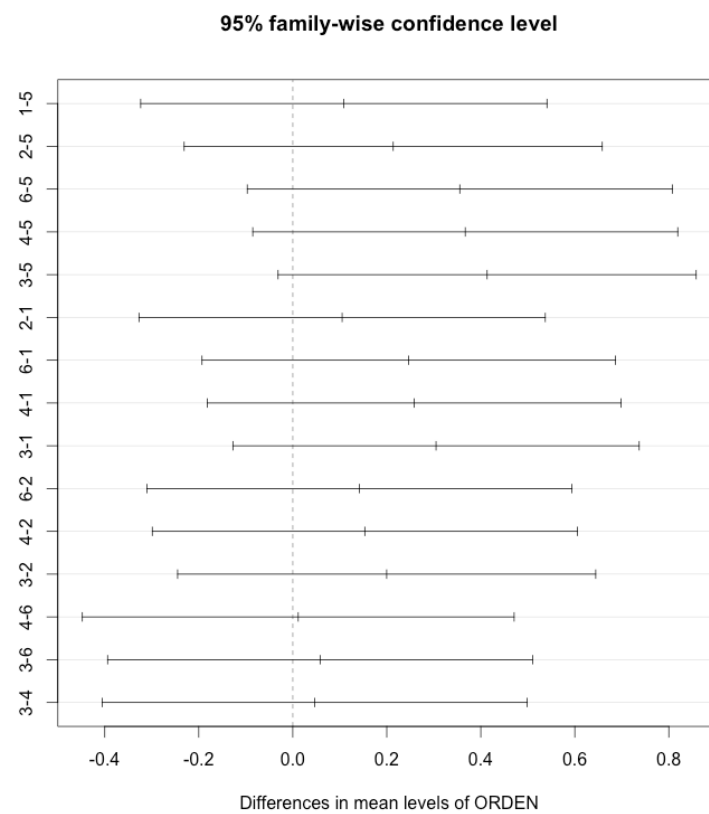


Figura 42: Test de Tukey para el factor Orden y la variable Tiempo MU-PRO

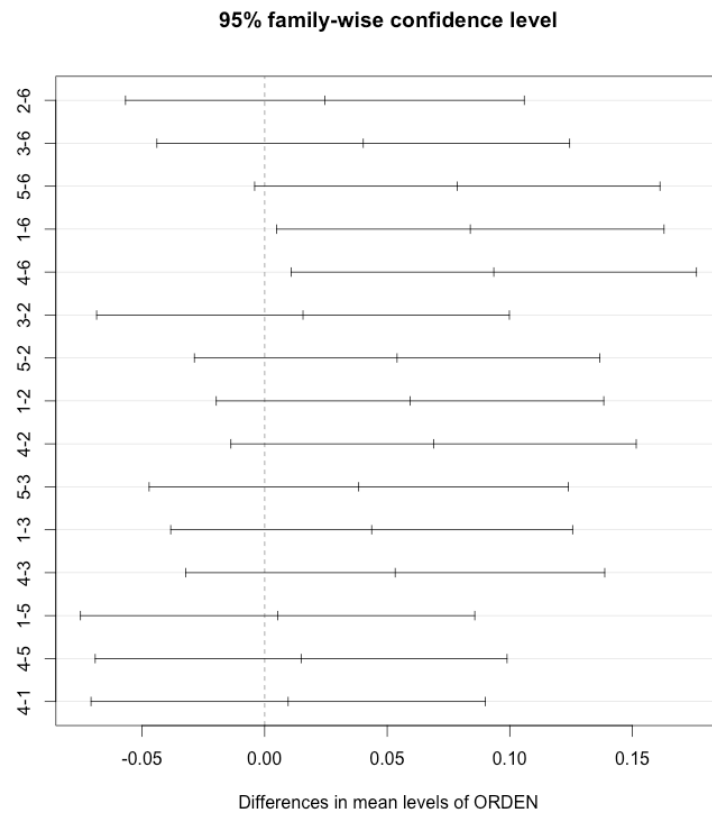


Figura 43: Test de Tukey para el factor Orden y la variable N° de Clicks MU-PRO

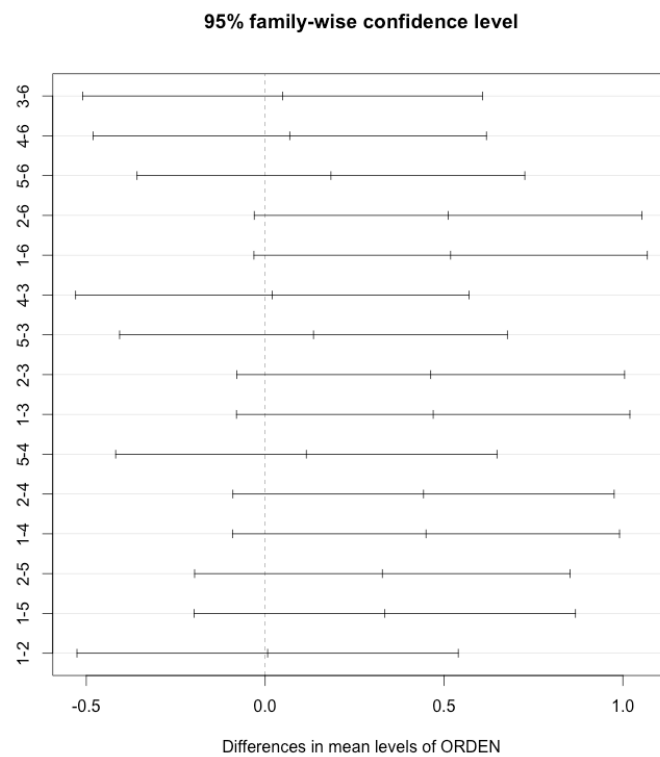


Figura 44: Test de Tukey para el factor Orden y la variable Tiempo MU-ABO

Anexo F. Cuestionario de Familiaridad

En este anexo se detalla el cuestionario de familiaridad presentado tanto en el experimento verdadero final como en el experimento piloto, éste se recoge en la Figura 45.

| | | | | |
|---|--|--|------------------------------------|--|
| Nombre o Alias:* | <input type="text"/> | | | |
| Sexo: * | <input type="radio"/> Masculino | <input type="radio"/> Femenino | | |
| Edad: * | <input type="radio"/> Menos de 18 | <input type="radio"/> 18-30 | <input type="radio"/> 31-40 | <input type="radio"/> 41-50 <input type="radio"/> 55 o más |
| Idiomas conocidos: * | <input type="checkbox"/> Castellano | <input type="checkbox"/> Inglés | <input type="checkbox"/> Francés | <input type="checkbox"/> Otros |
| ¿Eres profesional en el Área de la Informática ? | <input type="radio"/> No <input type="radio"/> Sí | | | |
| Profesión/Ocupación: | <input type="text"/> | | | |
| ¿Dónde usas internet con frecuencia?* | <input type="radio"/> En casa | <input type="radio"/> No uso internet | | |
| | <input type="radio"/> En el trabajo | <input type="radio"/> Otros | <input type="text" value="Otros"/> | |
| ¿Para qué usas internet?* | <input type="checkbox"/> Compras | <input type="checkbox"/> Educación | | |
| | <input type="checkbox"/> Trabajo | <input type="checkbox"/> Entretenimiento | <input type="text" value="Otros"/> | |
| ¿Qué aplicaciones usas con mayor frecuencia? | <input type="checkbox"/> Facebook | <input type="checkbox"/> Google | | |
| | <input type="checkbox"/> Twitter | <input type="checkbox"/> Yahoo | <input type="text" value="Otros"/> | |
| ¿Has realizado compras personalmente por internet?* | <input type="radio"/> Nunca | <input type="radio"/> Rara vez | <input type="radio"/> A veces | <input type="radio"/> Casi siempre <input type="radio"/> Siempre |
| ¿Has realizado compras por internet con ayuda de intermediarios?* | <input type="radio"/> Nunca | <input type="radio"/> Rara vez | <input type="radio"/> A veces | <input type="radio"/> Casi siempre <input type="radio"/> Siempre |
| ¿Qué tiendas en internet has usado para buscar o comprar artículos?* | <input type="checkbox"/> Amazon | <input type="checkbox"/> Otras Tiendas | | |
| | <input type="checkbox"/> eBay | <input type="checkbox"/> Ninguno | | |
| ¿Qué problemas percibes cuando compras o quieres comprar por internet?* | <input type="radio"/> Los artículos se presentan con precios inesperados <input type="radio"/> Las tiendas físicas ofrecen mejores precios y promociones <input type="radio"/> Siento inseguridad en el pago (tarjetas de crédito, etc.) <input type="radio"/> No encuentro problemas <input type="radio"/> Otros <input type="text"/> | | | |
| En líneas generales, ¿consideras una opción atractiva las compras por internet? | <input type="radio"/> No <input type="radio"/> Sí | | | |

Figura 45: Cuestionario de Familiaridad